UNIVERSIDAD AUTÓNOMA DE MADRID

FACULTAD DE CIENCIAS

DEPARTAMENTO DE BIOLOGÍA MOLECULAR

Unique roles of cohesin-SA2 in proliferation and gene regulation: impact on embryonic development

Magali M. A. De Koninck

Doctoral Thesis

Madrid, 2020

UNIVERSIDAD AUTÓNOMA DE MADRID

FACULTAD DE CIENCIAS

DEPARTAMENTO DE BIOLOGÍA MOLECULAR

Unique roles of cohesin-SA2 in proliferation and gene regulation: impact on embryonic development

Magali M. A. De Koninck

Graduate in Biotechnology

Thesis Director: Dr. Ana Losada

Centro Nacional de Investigaciones Oncológicas (CNIO)

Madrid, 2020

Funding

This thesis has been supported by and FPI fellowship awarded to Magali De Koninck (BES-2014-069166) and research grants from MINECO and FEDER Funds BFU2013-48481-R and BFU2016-79841-R.

Agradecimientos

En primer lugar, quiero agradecer a mi directora Ana Losada. Gracias por acogerme en tu laboratorio y confiarme este proyecto. Gracias por tener siempre tu puerta abierta para nosotros y motivarnos a trabajar dando ejemplo. Gracias por ser firme cuando lo necesitamos y comprensiva con nuestras frustraciones. Te agradezco todo lo que has hecho por mí, desde llevarme de paquete al primer *SMC meeting*, hasta ayudarme con cada una de mis presentaciones y reunirte conmigo semana sí, semana también. He aprendido mucho de ti y me siento muy afortunada de haber pasado por este laboratorio.

Gracias a mis compañeros de grupo por formar un gran equipo. A Ana Cuadrado, por tus ideas y ayuda. A Miguel, por enseñarme tanto (como tener paciencia con los MEFs y los FRAPs), tener siempre buena actitud y formar el dúo *Formaldehyde*. A las técnicos (Miry desde el principio, Marta R., Marta V. y Rocío) por todas vuestras horas de ayuda, aportáis muchísimo a nuestros proyectos. A Alek, por tu valiosa ayuda en materia de análisis genómicos. A Dani, por tomar su relevo y por tu mente brillante. A Sam, por tu amistad y *dinner club*, necesitaba a alguien como tú cuando llegué. A Carmen, mi gaditana, por el apoyo moral(es). A Dácil, por desprender tan buena energía.

Gracias a Juan Méndez y su laboratorio (Sara, Sergio, Karol, Marcos, Dani, Patri, Elena, Susana, Sara S.) por formar parte de algo que va más allá de los límites de un grupo de investigación. Gracias por las aportaciones en los *lab meetings* y por hacer que nuestro laboratorio tenga tan buen ambiente de trabajo. ¿No todo van a ser cohesinas, no?

Gracias a Paco Real y Eleonora Lapi por la colaboración que nació con un ratón de interés común antes de mi llegada. Gracias por nuestras reuniones conjuntas, espero que pronto podamos celebrar nuestro *paper*.

A Miguel Manzanares y Claudio Badía, gracias por acceder a formar parte del proyecto cuando estábamos bastante perdidos. Por vuestras ideas y por acogerme durante mi "estancia" para cacharrear en vuestro laboratorio. Claudio, me has enseñado muchísimo y te has volcado por completo en este proyecto, ¡gracias por darle un enorme impulso!

Gracias a todas las unidades del CNIO: Animalario, Ratones transgénicos, Citometría de flujo, Microscopía confocal, Genómica, Anticuerpos monoclonales. Vuestra labor ha sido crucial para este trabajo y valoro mucho vuestra ayuda. En especial gracias a Diego y los chicos de confocal por afrontar con tan buena actitud el desafío que suponen los FRAPs.

Gracias a mi comité de tesis: Marcos Malumbres, Paco Real y Alberto M. Pendás, por las reuniones e ideas.

Gracias a mis comienzos en la ciencia: a Mr. Galvin por transmitirme su pasión por la biología, y a Marta Artal y Trond Aasen por guiarme en mis primeros pasos en la poyata.

Gracias a las *Mutis*, habéis estado ahí desde nuestros primeros pinitos en la ciencia y en Sevilla. Cada una ha tomado un camino diferente pero al final siempre estáis ahí. Por más momentos con vosotras.

Gracias a Marian y las chicas de danza, por su positivismo y hacer que los lunes fueran menos lunes.

A las *Sestras*, gracias por compartir muchas horas, risas y lágrimas en el *Senaio*. Solo vosotros conseguís plasmar todas las frustraciones posibles en posturas de yoga. En el laboratorio se necesitan compañeros como vosotros. Gracias también por todos los momentos vividos fuera, desde el gallego hasta “lo alto” del Teide. En la vida se necesitan amigos como vosotros. Sois sin duda lo mejor que me ha dado esta tesis.

Gracias Paco, porque este camino no hubiera sido igual sin tenerte a mi lado. Gracias por aguantarme en momentos duros y por celebrar conmigo las pequeñas cosas de la vida. Gracias por los planes en diferido, los simulacros, los findes de visitas y los viajes que me han dado un respiro de la tesis. Me encanta tu visión del mundo y me ayudas a relativizarlo todo. Lo hemos conseguido. Ahora, ha llegado nuestro momento.

Gracias a mi familia por darme una vida plena y su apoyo incondicional. Gracias por los mensajitos positivos, las risas, el *haka*, las visitas, los *breakfasts at Tiffany’s*. Gracias por los findes de mimos: n*ergens beter dan thuis*! No me puedo imaginar una familia mejor. Gracias por siempre recordarme lo que realmente es importante.

¡Mil gracias a todos por ser un pedacito de esta tesis!

Table of contents

Table of contents

Summary

Resumen

Summary

Cohesin is a multiprotein complex composed of four subunits: SMC1, SMC3, Rad21 and SA. Within its ring-shape, two DNA fibers can be held together. Cohesin is best known for its role in cohesion by stably entrapping sister chromatids. In addition, cohesin can also form loops by bringing together distal regions from the same chromatid. This property establishes cohesin as a major organizer of interphase chromatin, which is essential for gene regulation among other processes. In somatic vertebrate cells, the SA subunit can be either SA1 or SA2, thus giving rise to two coexisting cohesin variants: cohesin-SA1 and cohesin-SA2. Despite their similarity, both proteins are functionally non-redundant, best exemplified by the lethality of SA1-null embryos. *STAG2*, the gene encoding SA2, is one of the most frequently mutated genes across multiple cancer types, and recently germline mutations have also been identified in developmental syndromes known as cohesinopathies.

In this Thesis, we generated a *Stag2* conditional knock out (cKO) mouse model to shed light onto the functional specificities of cohesin variants and address the consequences of SA2 loss both *in vitro*, in murine embryonic fibroblasts (MEFs), and *in vivo*, for embryonic development. We found that SA2 is dispensable *in vitro*, but its depletion leads to slower proliferation and reduced fidelity of chromosome segregation. We also show that the genome wide distribution of cohesin-SA2 in MEFs is not identical to that of cohesin-SA1. While both complexes can be found at sites also bound by the insulator protein CTCF, a fraction of cohesin-SA2 is present at enhancers independently of CTCF. These sites cannot be occupied by cohesin-SA1 even when SA2 is absent, which likely alters gene expression. We demonstrate that the association of cohesin-SA2 with chromatin is more dynamic than that of cohesin-SA1, a feature that likely explain the distinct distribution of the two variants and the different transcriptional signatures in MEFs lacking SA1 or SA2.

Despite being dispensable for cell viability *in vitro*, cohesin-SA2 is critical for embryonic development. SA2-null embryos die by mid-gestation, displaying a systemic developmental delay along with specific heart defects that result from a general reduction in proliferation and impaired morphogenesis resulting from altered transcription. We uncovered that a particular pool of cardiac progenitors requires SA2 to efficiently deploy into the elongating heart tube and contribute to its differentiation. These results support a unique role for cohesin-SA2 during embryogenesis that cannot be assumed by cohesin-SA1, and suggest a causal relationship between *STAG2* mutations and cardiac anomalies in cohesinopathy patients.

Resumen

La cohesina es un complejo multiproteico en forma de anillo compuesto por SMC1, SMC3, Rad21 y SA, que abraza la fibra de ADN. Además de su papel en el apareamiento de cromátidas hermanas o cohesión, puede formar lazos de cromatina al abrazar regiones distales de una misma cromátida. Esta propiedad establece a la cohesina como un importante organizador de la cromatina interfásica, esencial para la regulación génica entre otros procesos. En células somáticas de organismos vertebrados, la subunidad SA puede ser SA1 o SA2, dando lugar a dos variantes que coexisten en todas las células: cohesina-SA1 y cohesina-SA2. A pesar de su similitud, los dos complejos no son redundantes, siendo la mejor prueba de ello la letalidad de los embriones murinos carentes de SA1. *STAG2*, el gen que codifica SA2, es uno de los más mutados en cáncer, y recientemente se han identificado mutaciones en la línea germinal en pacientes de síndromes de desarrollo conocidos como cohesinopatías.

En esta Tesis hemos generado un modelo de ratón *knock out* condicional de *Stag2* para esclarecer la especificidad funcional de las variantes de cohesina y estudiar las consecuencias de eliminar SA2 tanto *in vitro*, en fibroblastos embrionarios de ratón, como *in vivo*, durante el desarrollo embrionario. Observamos que SA2 es dispensable *in vitro*, pero su depleción conlleva una proliferación más lenta y una menor fidelidad en segregación cromosómica. También mostramos que la distribución genómica de las dos variantes no es idéntica. Mientras ambas ocupan sitios donde también se une CTCF, una fracción de los complejos cohesina-SA2 se encuentra en *enhancers* independientemente de CTCF. Estos sitios no pueden ser ocupados por cohesina-SA1, incluso en ausencia de cohesina-SA2, lo que altera la expresión génica. Demostramos que la asociación de cohesina-SA2 a cromatina es más dinámica que la de cohesina-SA1, lo que probablemente determina su distribución genómica y contribuye a explicar que los cambios transcripcionales en células carentes de SA1 o SA2 sean muy diferentes.

A pesar de ser dispensable para la viabilidad celular in vitro, la cohesina-SA2 es crítica para el desarrollo embrionario. En su ausencia, los embriones mueren a mitad del período de gestación y presentan un retraso de crecimiento sistémico y defectos específicos en el corazón. Estos defectos se deben a una menor proliferación y a alteraciones transcripcionales que afectan a la migración de una población de progenitores cardiacos hacia el tubo cardiaco durante su elongación. Así pues, hemos identificado un papel único de la cohesina-SA2 en la morfogénesis del corazón en el desarrollo embrionario que sugiere una relación causal entre mutaciones en *STAG2* y las anomalías cardiacas de los pacientes de cohesinopatías.

Abbreviations

Abbreviations

AML	Acute myeloid leukemia
3C	Chromosome Conformation Capture
ASHF	Anterior secondary heart field
AVC	Atrioventricular canal
BrdU	Bromodeoxyuridine
CdLS	Cornelia de Lange Syndrome
ChIP-seq	Chromatin Immunoprecipitation coupled with deep sequencing
cKO	Conditional knock out
CNV	Copy number variation
co-IP	Co-Immunoprecipitation
CTCF	CCCTC-binding factor
DEG	Differentially Expressed Gene
Esco1/2	Establishment of Cohesion 1/2
FACS	Fluorescence Activated Cell Sorting
FC	Fold Change
FDR	False Discovery Rate
FHF	First heart field
FPKM	Fragments per kilobase per million of reads
FRAP	Fluorescence Recovery After Photobleaching
GO	Gene Ontology
GSEA	Gene Set Enrichment Analysis
H3P	Histone H3 phosphorylated at Serine 10
HC	Heart chambers
HDAC8	Histone Deacetylase 8
HEAT	Huntingtin, elongation factor 3 (EF3), protein phosphatase 2A (PP2A) and the yeast kinase TOR1
HR	Homologous Recombination
HSPC	Hematopoietic stem and progenitor cell
IB	Immunoblot
iFRAP	inverse Fluorescence Recovery After Photobleaching
IFT	Inflow tract
ISL1	Islet1 transcription factor
KO	knock out
LA	Left atrium
LV	Left ventricle
MEFs	Mouse Embryonic Fibroblasts
mESC	Mouse Embryonic Stem Cell
NES	Normalized Enrichment Score
NIPBL	Nipped-B-like protein
ns	non-significant
NT	Neural tube
OFT	Outflow tract
Pds5	Precocious dissociation of sisters 5
PP2A	Protein Phosphatase 2A
PSHF	Posterior secondary heart field

RA	Right atrium
RAD21	Double-strand-break repair protein 21 homolog
RBS	Roberts Syndrome
RNA-seq	RNA-sequencing
RNApolII	RNA polymerase II
RT-qPCR	Retrotranscription followed by quantitative PCR
RV	Right ventricle
SA	Stromal Antigen (protein)
scRNA-seq	Single cell RNA-seq
SEC	Super-elongation complex
SEM	Standard error of the mean
Sgo1	Shugoshin 1
sgRNA	short guide RNA
SHF	Secondary heart field
siRNA	small interfering RNA
SMC	Structural Maintenance of Chromosomes
STAG	Stromal Antigen (gene)
TAD	Topologically Associating Domain
tSNE	t-Distributed Stochastic Neighbor Embedding
TSS	Transcription Start Site
TUNEL	Terminal deoxynucleotidyl transferase dUTP nick end labeling
WAPL	Wings apart-like protein
WT	wild type
XCi	X chromosome inactivation

Introduction

Introduction

1. The cohesin complex

Cohesin is a large multiprotein complex conserved from yeast to human, which orchestrates multiple DNA transactions within the nucleus. Cohesin is one of the complexes of the Structural Maintenance of Chromosomes (SMC) family that exist in eukaryotic cells, together with condensin and the SMC5/6 complex, but SMC complexes exist in all kingdoms of life (Uhlmann 2016). Cohesin subunits were originally identified in yeast screenings as mutants of sister chromatid cohesion (Guacci *et al.* 1997; Michaelis *et al.* 1997) and were soon after shown to form a complex in *Xenopus laevis* egg extracts and human cells (Losada *et al.* 1998, 2000; Sumara *et al.* 2000). Since then, cohesin has been attributed a well-established role in mediating sister chromatid cohesion, which is essential for chromosome segregation in mitosis and meiosis and DNA repair by homologous recombination (HR). Emerging research over the last decade has also recognized cohesin as major player in higher-order genome organization through chromatin looping, thereby regulating gene expression and replication timing (Nasmyth & Haering 2009). Additional roles of cohesin include fork progression, architecture of DNA replication factories, VDJ recombination and locus rearrangement, most of which are the consequence of genome organization.

1.1 Cohesin composition

Cohesin is a ring-shaped complex composed of two SMC proteins, a kleisin subunit and a Huntingtin, elongation factor 3 (EF3), protein phosphatase 2A (PP2A) and yeast kinase TOR1 (HEAT)-repeat protein (Fig. I1A). SMCs are long rod-like proteins that fold on themselves to create a hinge domain on one end and an ATPase head domain containing the N- and C-terminus on the other, separated by long anti-parallel coiled coils. When SMCs heterodimerize they form a symmetrical structure that interacts through the hinge domains. The non-SMC subunits form a globular domain in between the catalytic domains (Anderson *et al.* 2002) (Fig. I1B).

The SMC proteins that compose cohesin are SMC1 and SMC3. A tripartite ring is formed by interaction with the kleisin Rad21, that bridges the head domains through interaction of its C- and N-terminal domains near the SMC1 and SMC3 heads, respectively. The HEAT-repeat protein Stromal Antigen (SA) is the fourth subunit that completes the core

cohesin complex through binding to Rad21 (Gligoris *et al.* 2014; Hara *et al.* 2014). This region serves as an interaction platform for cohesin regulatory proteins.

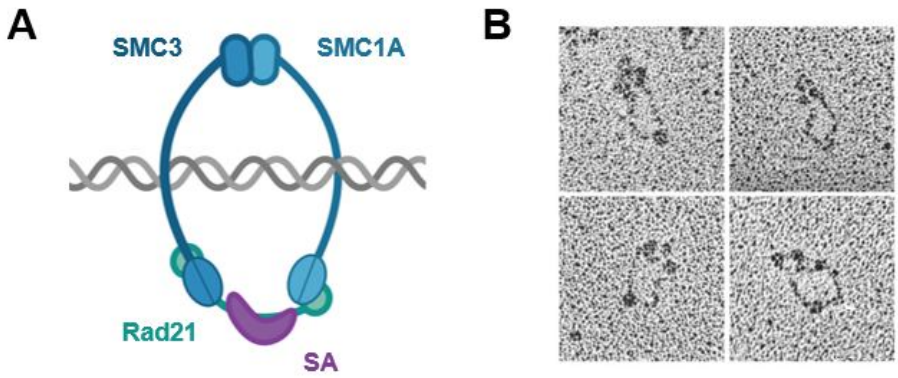


Figure I1. The cohesin complex.

A. Composition of the somatic cohesin complex. SMC3 and SMC1A form a heterodimer tethered by the hinge domains. SMC head domains are bridged by the kleisin subunit Rad21 to form a tripartite ring that is completed by binding of the HEAT-repeat protein Stromal Antigen (SA). In somatic vertebrate cells, the SA subunit can be either SA1 or SA2. Meiosis-specific versions exist for all cohesin subunits except SMC3 (not depicted).
B. Electron micrographs of human cohesin complexes purified from HeLa cells. Taken from Anderson *et al.* 2002.

Intriguingly, two versions of cohesin coexist in somatic vertebrate cells, that contain one of two variants of the SA subunit: SA1 or SA2. This Thesis focuses only on somatic cohesin complexes, although meiosis-specific versions exist for all cohesin subunits except SMC3 (Biswas *et al.* 2016; Llano *et al.* 2012). This leads to a wide array of different cohesin complexes depending on the combination of subunits.

1.2 DNA entrapment by cohesin

Evidence from a number of *in vitro* and *in vivo* studies supports a model in which cohesin can topologically entrap a chromatin fiber within the lumen of its ring-shaped structure (Gruber *et al.* 2003; Haering *et al.* 2008). DNA entrapment is essential for cohesin functions. Both for its role in cohesion between sister chromatids (*trans* entrapment, Fig. I2A) and for its role in chromatin looping (*cis* entrapment, Fig. I2B), cohesin would need to tether two different DNA strands at the same time.

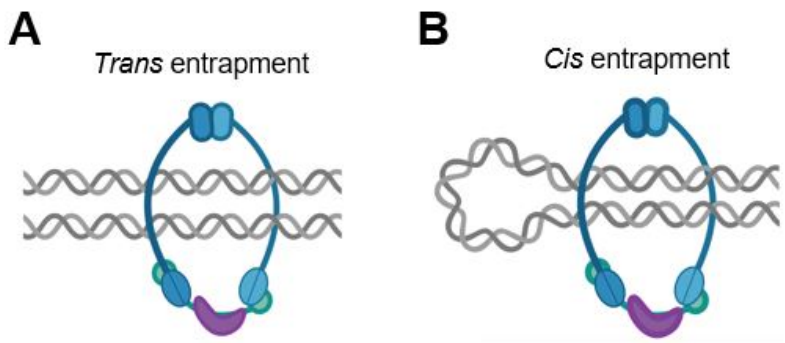


Figure I2. Two modes of cohesin-mediated entrapment of chromatin fibers.
Cohesin can embrace (A) two sister chromatids in *trans* to perform cohesion or also B) entrap two distal DNA fragments of the same chromatid in *cis* to form loops involved in chromatin organization.

Different models have been proposed to account for this (Fig. I3). The classical "embrace" model suggests that both chromatin fibers are entrapped within the lumen of one cohesin ring (Gruber *et al.* 2003; Haering *et al.* 2008). A second model, known as the "handcuff" model, proposes that two cohesin rings associated through their SA subunit each entrap one chromatin fiber (Zhang *et al.* 2008a). Alternatively, one cohesin complex could entrap two fibers within two potential compartments in the same ring in a "two gates" model that could reconcile the two previous ones. This idea is driven by the rapidly growing amount of data about the cohesin ring architecture (Nishiyama 2019). Whether the two tethered DNA strands are embraced by a single cohesin complex or two interacting complexes is still a matter of controversy and could depend on the functional nature of the tethering. Very recent data has shown that, at least *in vitro*, a single cohesin is responsible for *cis* loop formation (Davidson *et al.* 2019).

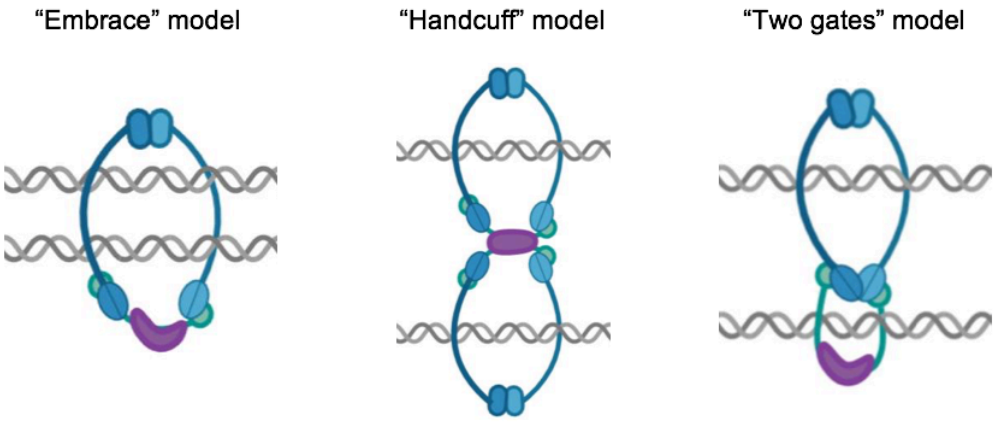


Figure I3. Current models of cohesin-mediated entrapment of chromatin fibers.
The classical "embrace" model suggests that two fibers are held together within the single lumen of the cohesin ring. The "handcuff" model proposes that each fiber is held by different cohesin complexes that interact. The "two gates" model suggests that SMC head tethering creates two distinct compartments within one cohesin ring, each entrapping a fiber.

2. Dynamic association of cohesin with chromatin throughout the cell cycle

Cohesin's association with DNA during the different phases of the cell cycle is highly dynamic and tightly regulated by different processes: cohesin loading and unloading, cohesion establishment and cohesion dissolution. Fluorescence Recovery After Photobleaching (FRAP) studies have shed light on the distinct cohesin populations in the cell. While a fraction of cohesin remains soluble in the nucleus, another is dynamically associated to chromatin throughout the cell cycle. In addition, a third and very stable cohesin population arises from S phase onwards to mediate sister chromatid cohesion (Gerlich *et al.* 2006).

2.1 Loading and unloading of cohesin during interphase

In vertebrate cells, cohesin is loaded onto chromatin as cells exit mitosis, in late telophase and early G1. The loading reaction is mediated by the heterodimer formed by Nipped B-like protein (NIPBL) and Mau2, that binds to cohesin and stimulates ATP hydrolysis by the SMC heads (Arumugam *et al.* 2006; Gillespie & Hirano 2004). This allows entrapment of DNA through opening of the "entry gate", consisting of separation of the SMC hinge domains (Gruber *et al.* 2006) (Fig. I4).

Once loaded, cohesin binding to chromatin is dynamic. Two other HEAT-repeat proteins associate with cohesin to modulate its association to chromatin: Precocious dissociation of sisters (Pds5), which in somatic vertebrate cells can be Pds5A or Pds5B, and Wings-apart like protein (WAPL) (Gandhi *et al.* 2006; Kueng *et al.* 2006; Losada *et al.* 2005; Sumara *et al.* 2000). Both proteins work together to promote cohesin unloading through the ATP-dependent opening of an "exit gate" consisting in separation of the SMC3-Rad21 interface (Gligoris *et al.* 2014; Huis In't Veld *et al.* 2014; Murayama & Uhlmann 2015; Ouyang *et al.* 2016) (Fig. I4).

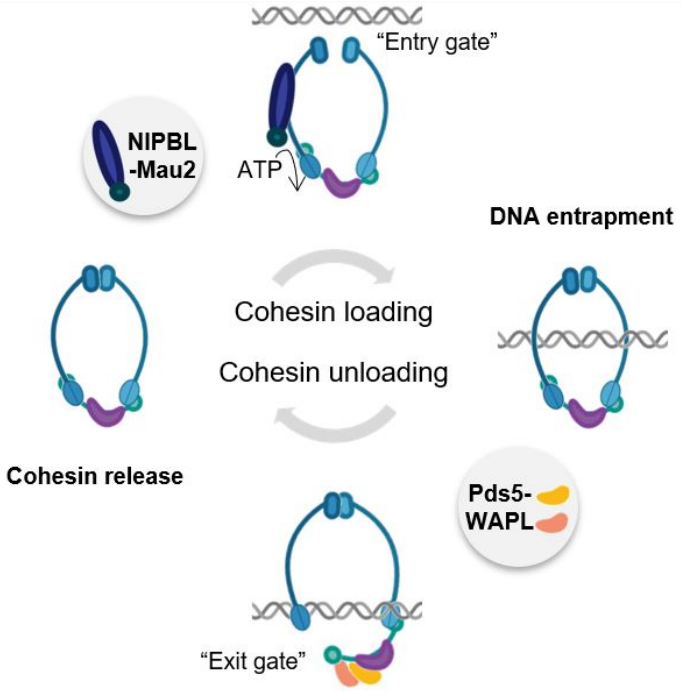


Figure I4. Mechanism for dynamic loading and unloading of cohesin.
Cohesin is loaded onto chromatin in interphase by the cohesin loader complex formed by the NIPBL-Mau2 heterodimer in an ATP-dependent manner. This allows entrapment of a chromatin fiber by stimulating the opening of the "entry gate" between the SMC hinges. Once loaded, cohesin is susceptible to removal mediated by Pds5 and WAPL. Binding of both factors promotes opening of the "exit gate", the interface between Rad21 and SMC3.

2.2 Cohesion establishment after DNA replication

From the moment they emerge from the replication fork, replicated chromatids need to be held together until their separation during mitosis. This process is known as cohesion and promotes faithful DNA repair by homologous recombination (HR) in interphase and ensures accurate chromosome segregation. During DNA replication a fraction of the cellular cohesin pool becomes stably bound to chromatin (Gerlich *et al.* 2006) and is responsible for this tethering, in a tightly regulated manner.

Cohesion establishment requires two events: SMC3 acetylation and Sororin binding (Fig. I5). Both events must occur in the context of DNA replication (Ladurner *et al.* 2016; Lafont *et al.* 2010). In vertebrates, SMC3 acetylation is carried out at residues K105 and K106 of the head domain by two cohesin acetyltransferases, Esco1 and Esco2. The acetylated cohesin complex can then be bound by Sororin, which displaces WAPL and counteracts its unloading activity (Ladurner *et al.* 2016; Nishiyama *et al.* 2010). Both events are facilitated by Pds5. Acetylated cohesin complexes bound by Sororin are the ones responsible for cohesion between sister chromatids and persist until mitosis.

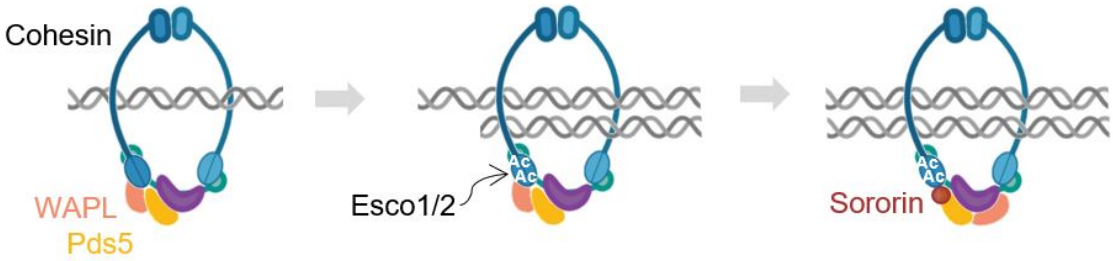


Figure I5. Cohesion establishment requires SMC3 acetylation and Sororin binding. Once DNA is replicated, the cohesin acetyltransferases Esco1/Esco2 acetylate two residues in the head domain of SMC3. Next, Sororin is recruited through Pds5, which causes WAPL displacement. Acetylated and Sororin-bound cohesin is now stably bound to chromatin, promoting cohesion between newly replicated sister chromatids.

2.3 Cohesion dissolution in mitosis

Upon mitotic entry, cohesin undergoes a stepwise dissociation from chromatin, first in prophase and later in anaphase (Waizenegger *et al.* 2000) (Fig. I6). Most cohesin is removed in a first step regulated by phosphorylation, known as the prophase pathway. Phosphorylation of SA by Polo-like kinase 1 (Plk1) and release of Sororin upon phosphorylation by cyclin-dependent kinase 1 (CDK1) and AuroraB restore WAPL activity (Gandhi *et al.* 2006; Hauf *et al.* 2005; Losada *et al.* 2002; Nishiyama *et al.* 2013; Shintomi & Hirano 2009). Cohesin can then be released by opening of the cohesin "exit gate". However, a small remaining fraction of cohesin enriched around centromeres resists the prophase dissociation pathway to keep sister chromatids together until anaphase, through different mechanisms. Shugoshin 1 (Sgo1) and protein phosphatase 2A (PP2A) antagonize the phosphorylation of SA and Sororin (Liu *et al.* 2013), while Sgo1 by itself competes with WAPL for binding to cohesin (Hara *et al.* 2014). The inhibitory chaperone Securin also avoids cleavage of Rad21 by Separase (Hauf *et al.* 2001; Lin *et al.* 2016; Uhlmann *et al.* 1999).

The cell enters anaphase when all chromosomes are correctly bioriented, that is, with sister kinetochores attached to microtubules from opposite spindle poles. This requires dissociation of all remaining cohesin, this time by a cleavage mechanism. Among other events, Securin is targeted for proteosomal degradation by the Anaphase Promoting Complex/Cyclosome (APC/C), which triggers Rad21 cleavage by Separase and allows sister chromatid separation (Musacchio & Salmon 2007). In late telophase and early G1, cohesin can again populate chromatin after deacetylation of SMC3 by HDAC8 (Deardorff *et al.* 2012) and the cell cycle can start again.

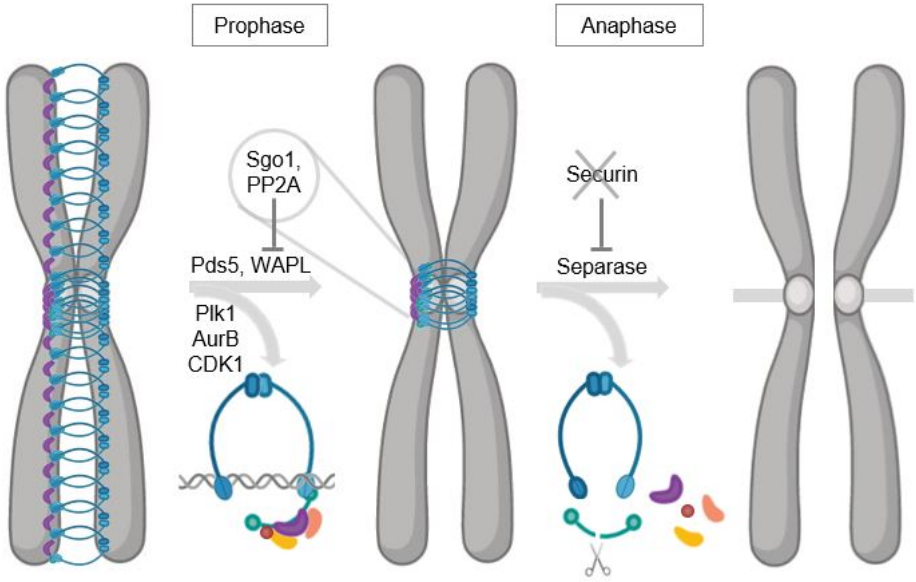


Figure I6. Different mechanisms promote stepwise removal of cohesin during mitosis. Most cohesin is unloaded in prophase through Pds5-WAPL-dependent dissociation, as in Fig. I3, mediated by phosphorylation signals. Cohesin at the centromeres is protected by the activity of Shugoshin 1 (Sgo1) and protein phosphatase (PP2A) to keep chromatids together until they are correctly bioriented. Cleavage-mediated removal of cohesin by Separase allows anaphase entry. See text for more details.

3. Cohesin in genome organization

In eukaryotes, the genome does not exist as a linear molecule. Imaging-based methods combined with chromosome conformation capture (3C) technologies with increasing resolution have shed light on the complex genomic organization. DNA is hierarchically packaged inside the nucleus, forming multiscale structural units of increasing complexity: chromatin fibers, chromatin loops, topologically associating domains (TADs), compartments and chromosome territories (Bonev & Cavalli 2016) (Fig. I7). Packaging provides on one side compaction, but also functional compartmentalization on all levels that modulates nuclear processes such as cell division, DNA replication and gene expression. Overall, eukaryotic genome architecture is crucial for cell fate and animal development (Zheng & Xie 2019).

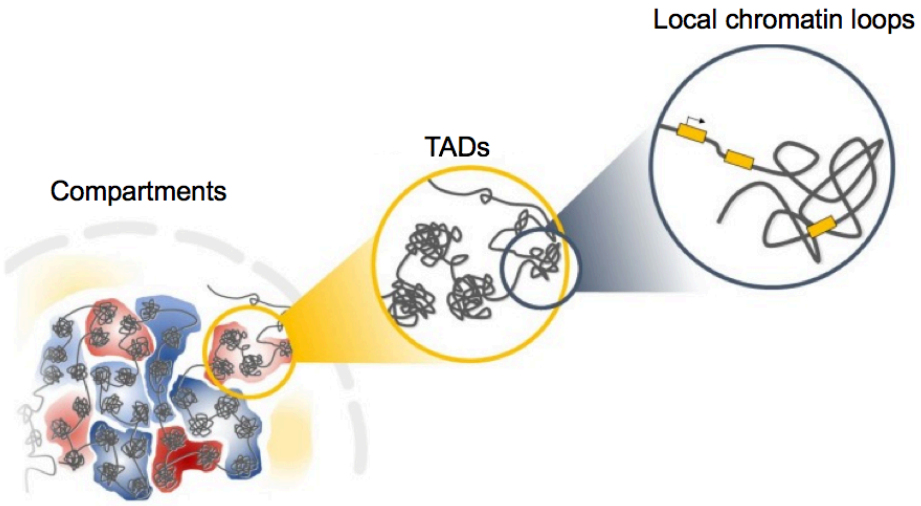


Figure I7. Hierarchical chromatin organization at different scales in eukaryotic genomes.
Chromosomes are segmented into active and inactive compartments. Within, topologically associating domains (TADs) form functional blocks. Inside, smaller intra-TAD chromatin loops are established. Taken from https://genominfo.org/.

Multiple components regulate genomic architecture, such as transcription factors, non-coding RNAs (ncRNAs) and architectural proteins. Among them, cohesin and CCCTC-binding factor (CTCF) play a very prominent role in mediating dynamic and functional interactions in three dimensions.

Initial studies using chromatin immunoprecipitation followed by deep sequencing (ChIP-seq), to assess the genome-wide distribution of cohesin, revealed an extensive colocalization of cohesin with the architectural protein CTCF (Parelho *et al.* 2008; Rubio *et al.* 2008; Wendt *et al.* 2008a). This zinc-finger protein binds DNA directly and early work recognized its role in enhancer/promoter insulation (Bell & Felsenfeld 2000; Hark *et al.* 2000). Later, chromosome conformation studies showed the importance of cohesin and CTCF in forming chromatin contacts and overall chromosome organization (Sofueva *et al.* 2013; Zuin *et al.* 2014).

TADs are often demarcated by cohesin and CTCF and are thought to regulate transcription by facilitating interactions between enhancers and promoters present in the same TAD, while restricting interactions with regulatory elements from different TADs. Genes within the same TAD tend to have similar gene expression dynamics, suggesting their role in coordinating the activity of a group of genes (Gibcus & Dekker 2013).
An attractive model for TAD generation proposes that, once loaded, cohesin can extrude DNA to generate progressively longer chromatid loops until it dissociates by WAPL-mediated release or until it gets stalled by an obstacle, such as CTCF binding (de Wit *et*

al. 2015; Fudenberg *et al.* 2016; Haarhuis *et al.* 2017; Sanborn *et al.* 2015) (Fig. I8). A very recent study indeed confirmed that cohesin can act as a loop extruder machine *in vitro* and forms progressive chromatin loops in a manner that depends on its ATPase activity, stimulated by the cohesin loader (Davidson *et al.* 2019).

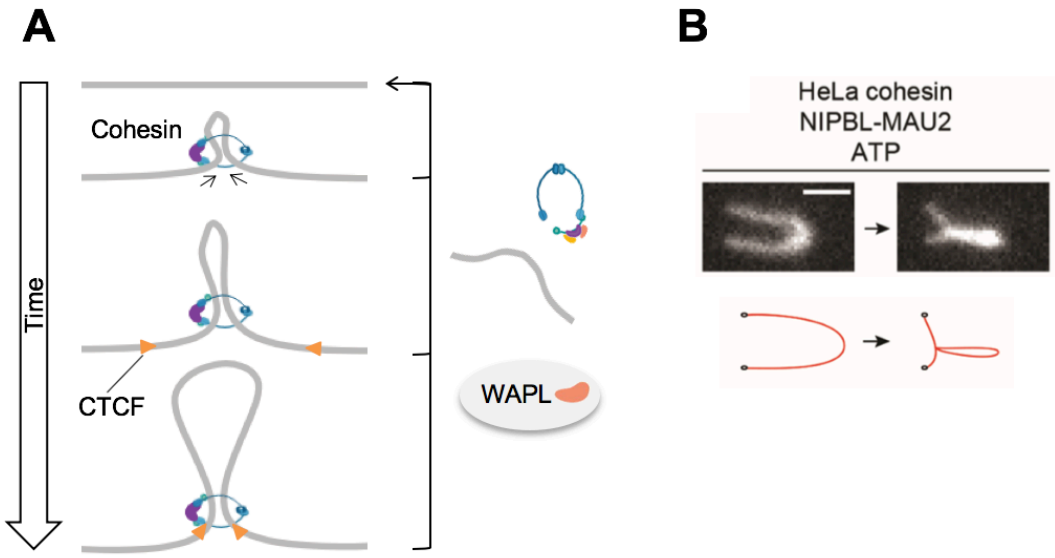


Figure I8. Model for cohesin as a loop extrusion factor.

A. The loop extrusion model for genome folding can explain how certain regions of chromosomes stay close together and why cohesin often colocalizes with CTCF at its binding sites, forming the anchor of a chromatin loop. According to this model, a loop is fed in a progressive manner through the cohesin ring, until cohesin encounters CTCF or until it is released from chromatin by WAPL activity.
B. *In vitro* evidence for a direct role of cohesin in loop extrusion. Biochemically reconstituted human cohesin complexes form loops in DNA tethered at both ends to a flow chamber, in the presence of the cohesin loader and ATP. Taken from Davidson *et al.* 2019.

The loop extrusion model could explain why cohesin and CTCF colocalize and why increasing cohesin residence time on chromatin causes formation of longer loops (Haarhuis *et al.* 2017; Wutz *et al.* 2017). It is also consistent with the fact that there is very little colocalization of cohesin with its loader (Busslinger *et al.* 2017; Wendt *et al.* 2008b), likely due to sliding of cohesin away from the loading site. This adds a new dynamic component to the complex, beyond loading and unloading. Why cohesin stops at CTCF sites remains unclear, since CTCF binds to chromatin much more dynamically than cohesin (Hansen *et al.* 2017). Some evidences support a role for Pds5 in this arrest (Wutz *et al.* 2017), maybe mediated by cohesin acetylation, which in G1 is mediated by Esco1 exclusively and requires Pds5 (Minamino *et al.* 2015).

CTCF-cohesin sites can also be found within TADs and contribute to cell-type specific sub-TAD organization (Phillips-Cremins *et al.* 2013). In addition, CTCF-independent

cohesin sites have also been identified, in which the complex occupies regions bound by tissue-specific transcription factors or transcriptional regulators such as Mediator (Faure *et al.* 2012; Kagey *et al.* 2010; Schmidt *et al.* 2010). These findings affirm the role of cohesin in establishing contacts at multiple scales.

4. Two non-redundant versions of cohesin in somatic cells

Somatic vertebrate cells express two paralogs of the cohesin SA subunit, SA1 and SA2 (Losada *et al.* 2000; Sumara *et al.* 2000). Their presence in the ring is mutually exclusive, giving rise to two distinct cohesin variants: cohesin-SA1 and cohesin-SA2 (Fig. I9A).

4.1 Sequence and structure of SA1 and SA2

SA1 and SA2 are proteins of 1258 and 1231 aminoacids, respectively. Both SA proteins contain 17 HEAT repeats (Hara *et al.* 2014). They have a high overall degree of sequence homology along the central region but differ in their N- and C-terminal regions (Fig. I9B). At its N-terminus, SA1 contains an AT-hook that has been proposed to mediate association to telomeric DNA (Bisht *et al.* 2013). This motif is absent in SA2.

Interestingly, crystallization attempts of the full-length SA subunits have failed, as the two studies that have solved the crystal structure lack the protein termini. In one, the yeast ortholog Scc3 was purified from two species of yeast but was only able to crystallize with trimmed N- and C-terminal extensions of the protein (Roig *et al.* 2014). In the other, human SA2 was crystallized as part of a cohesin RAD21-SA2 subcomplex containing SA2 residues 80-1260 (Hara *et al.* 2014) (Fig. I9C). Both studies agree that the terminal regions are unstructured and therefore impede crystallization. However, these regions are the most variable between SA homologs and may hold the key to their specificities.

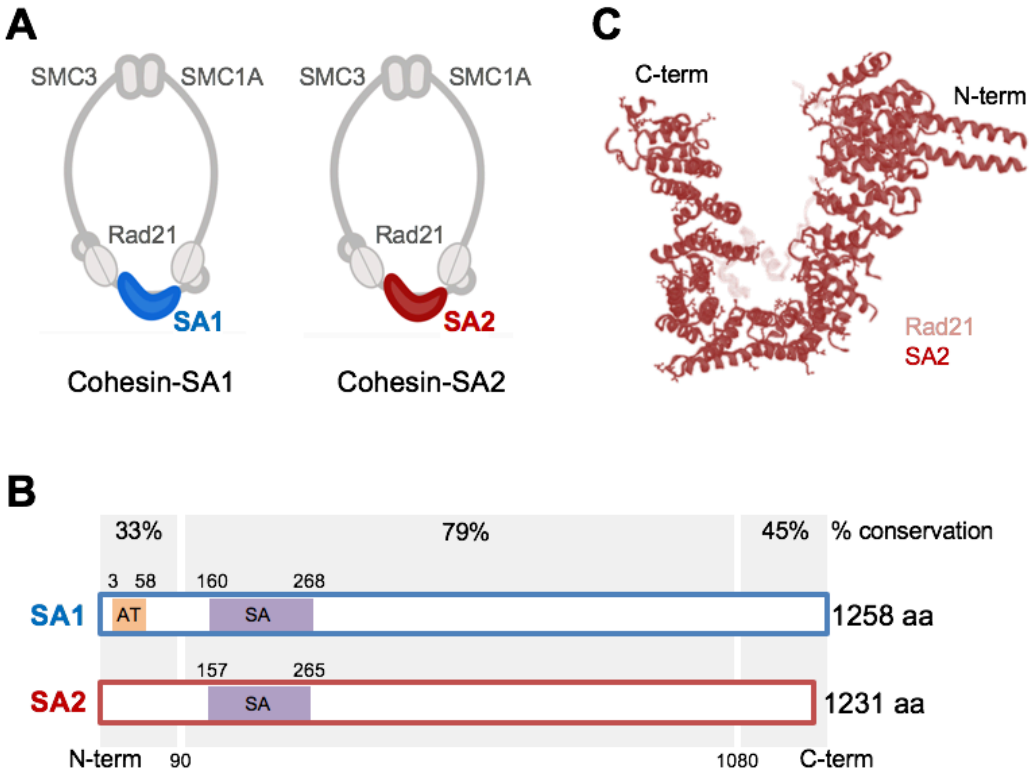


Figure I9. Two different cohesin complexes coexist in somatic vertebrate cells.

A. Cohesin complexes carrying either SA1 or SA2 are depicted.
B. Sequence conservation represents the percentage of amino acids (aa) that are identical between given regions. Conservation is highest in the central region (79% between aa 91-1080), whereas the N and C terminal ends are much more variable (33% conservation from aa 1-90 and 45% from aa 1080-end, respectively).
C. Dragon-shaped crystal structure for SA2 (dark red) containing residues 80-1260 together with the Rad21 interface (light red). Adapted from Hara *et al.* 2014.

4.2 Unique roles of cohesin-SA1 and cohesin-SA2

Despite their sequence similarity and their comparable chromatin-association dynamics throughout the cell cycle, cohesin-SA1 and cohesin-SA2 have non-redundant functions. Analysis of mouse and human cells depleted of SA1 or SA2 revealed a preferential contribution towards telomeric and centromeric cohesion, respectively (Canudas & Smith 2009; Remeseiro *et al.* 2012a) (Fig. I10). In mouse embryonic fibroblasts (MEFs) lacking SA1, telomeric cohesion defects lead to defects in telomere replication and, in turn, segregation defects (Remeseiro *et al.* 2012a). Other studies have defined preferential roles of cohesin-SA2 in other processes, such as recruitment to UV-induced double-strand breaks (DSBs) (Kong *et al.* 2013), repression of transcription near DSBs to allow correct repair (Meisenberg *et al.* 2019) and replication fork progression (Mondal *et al.* 2019).

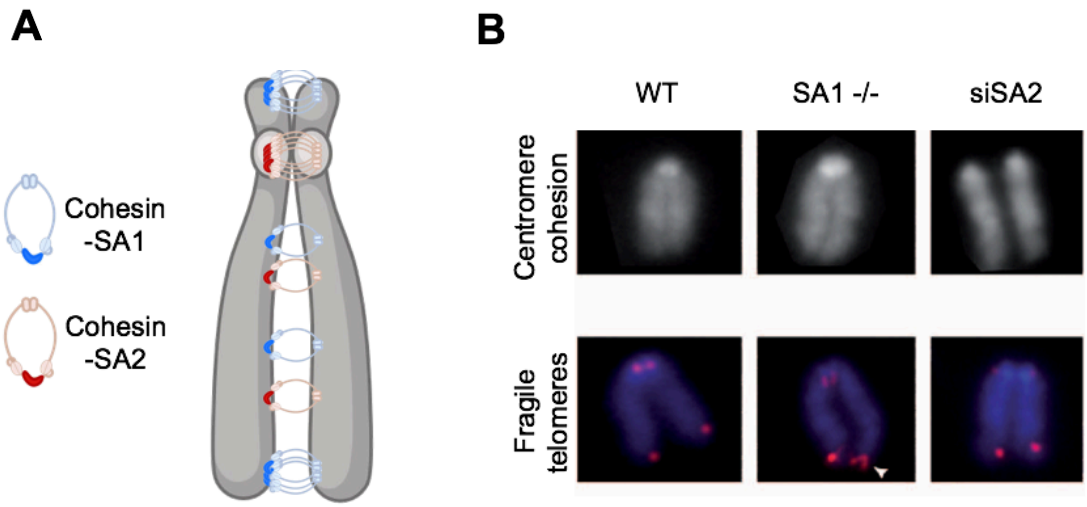


Figure I10. Specific roles of cohesin variants in cohesion.

A. Schematic representation of a specific contribution of SA1 to telomere cohesion and SA2 to centromere cohesion. Both contribute to cohesion along chromosome arms.
B. Representative images of metaphase chromosomes from MEFs wild type, SA1-null (SA1-/-) and depleted of SA2 (siSA2). Top, stained with DAPI (grey), showcasing centromere cohesion defect in the absence of SA2. Bottom, stained with DAPI (blue) and a telomeric repeat FISH probe (red), showing irregular or "fragile" telomeres that reflect defective telomere replication in the absence of SA1. Taken from Cuadrado *et al.* 2012.

On the other hand, the vast majority of studies related to the role of cohesin in chromatin organization have done so from the perspective of one of the core cohesin components (SMC1, SMC3 or Rad21). Global cohesin depletion or removal of cohesin from chromatin by NIPBL depletion leads to an overall loss of TADs and loops (Rao *et al.* 2017; Schwarzer *et al.* 2017).

In these studies, however, differences between SA1 and SA2 have not been addressed. During the course of this Thesis, our group has published two studies within this scope that have uncovered unique functions of cohesin-SA1 and cohesin-SA2 in genome organization. In human cell lines, both cohesin variants were both found at sites also bound by CTCF, but cohesin-SA1 played a stronger role in stabilizing and demarcating TAD boundaries. In addition, cohesin-SA2 also occupied sites without CTCF, involved in more local contacts that regulate tissue-specific transcription. At these sites, cohesin-SA2 could not be replaced by cohesin-SA1 (Kojic *et al.* 2018). In mouse embryonic stem cells (mESCs), a specific contribution of cohesin-SA2 was described to silencing of lineage-commitment genes through Polycomb recruitment and activation of pluripotency genes through local contacts involving superenhancers (Cuadrado *et al.* 2019).

Despite these differences, either cohesin is sufficient to allow cell proliferation and maintain viability in cultured cells (Mondal *et al.* 2019; Remeseiro *et al.* 2012a; van der Lelij *et al.* 2017). During embryonic development, however, cohesin-SA1 is essential, suggesting that there are functions that cohesin-SA2 is unable to fulfill. SA1-deficient mouse embryos from a constitutive knock out mouse model previously generated in our group are embryonic lethal starting at E11.5, although some embryos exceptionally survive until later stages of development (E17.5, Fig. I11). When analyzed, they displayed severe developmental abnormalities and hypoplasia at the organismal level, as well as gene deregulation at the cellular level (Remeseiro *et al.* 2012b).

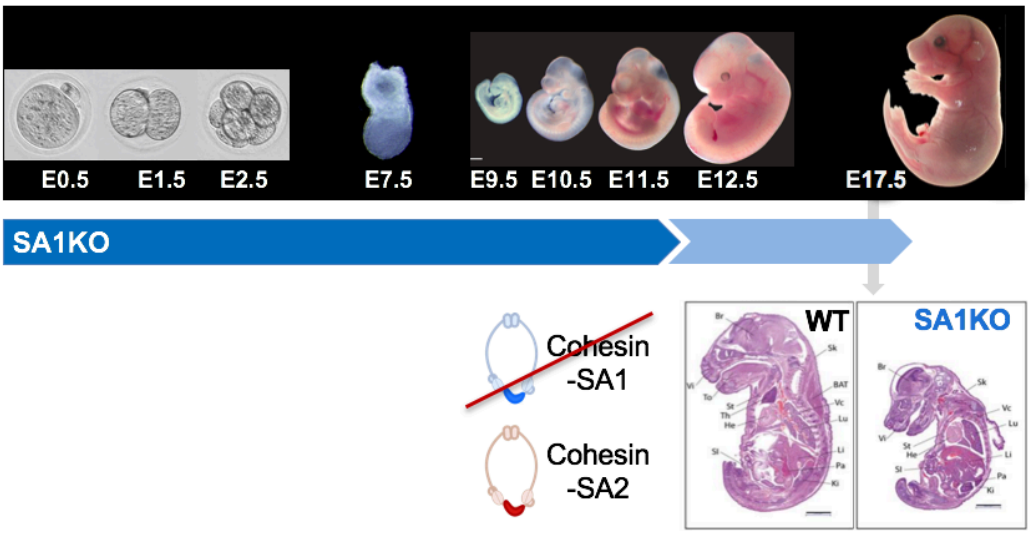


Figure I11. SA1-null mice are embryonic lethal.
Lethality of constitutive SA1KO embryos onsets at E11.5 but some survive until later stages. Wild type (WT) and SA1 knock out (SA1KO) embryos at E17.5 are shown. Data from Remeseiro *et al.* 2012b.

Overall, while cohesin-SA1 and cohesin-SA2 may have roles in which both complexes are interchangeable, it has become increasingly clear that they have non-redundant functions. Further elucidating these functions is one of the main goals of this Thesis and can hold relevant implications for development and disease, as highlighted in the next section.

5. Cohesin in development and disease

5.1 Cohesin and development

The development of a complex multicellular organism requires two fundamental processes: one is cell proliferation, needed for growth; the other is cell differentiation, in order to generate specialized cells and structures. Cohesin plays key roles in both these processes and is therefore essential for coordinating development (Remeseiro *et al.* 2013b). Several mutant mouse models for cohesin have been reported. Homozygous mutations in common cohesin subunits Rad21 or SMC3 are early embryonic lethal, but heterozygous mutants can be tolerated. Rad21 heterozygous mice display increased radiosensitivity and defective DSB repair (Xu *et al.* 2010), while SMC3 heterozygous mice show reduced body weight, craniofacial defects and impaired hematopoiesis (Wang *et al.* 2019; White *et al.* 2013). Mouse models with tissue-specific deletion of cohesin have also been reported. Depletion of SMC3 in the hematopoietic compartment leads to aberrant hematopoiesis with an altered balance between self-renewal and differentiation (Viny *et al.* 2015), while depletion in the brain leads to an abnormal neurological and behavioral phenotype (Fujita *et al.* 2017).

Given the wide scope of cohesin functions, its dysfunction is likely to alter many biological processes simultaneously, but not all of them have proven to be equally sensitive. While sufficient cohesion can be achieved with low cohesin levels (Carvalhal *et al.* 2018; Heidinger-Pauli *et al.* 2010), developmental defects and transcriptional alterations can arise with only small reductions (e.g. NIPBL heterozygous mice; Kawauchi *et al.* 2009; Muto *et al.* 2011). It is conceivable that cohesin's role in gene expression, for instance, may be much more sensitive to cohesin dosage than cohesion-related functions.

Three-dimensional genome architecture modulates biological processes including DNA replication, cell division and transcription, and as such is crucial for cell differentiation and multicellular organism development (Zheng & Xie 2019). As explained before, cohesin and CTCF have a crucial role in genome organization and mediate long-distance and local contacts that orchestrate gene regulation. They contribute to the genomic landscape that is established very early during embryonic development and are involved in later contact rewiring in the transition from a totipotent to a lineage-committed state (Hug & Vaquerizas 2018), ultimately affecting development and cell fate. It is not surprising that defects in regulatory contacts mediated by cohesin and CTCF can have

pathogenic consequences and lead to developmental abnormalities (Lupiáñez *et al.* 2016). Studies have shown that TAD disruption can cause “enhancer adoption” or establishment of ectopic enhancer-promoter contacts across regions that would normally be insulated, causing misexpression and disease (Ibn-Salem *et al.* 2014; Lupiáñez *et al.* 2015). TAD reorganization is also expected to play a role in the pathogenesis of cancer, for example in acute myeloid leukemia (AML) (Gröschel *et al.* 2014). Aberrant expression of single genes by altered local intra-TAD contact can also have potentially pathogenic consequences.

Cohesin pathway mutations lie at the origin of developmental syndromes, discussed below, highlighting its importance for correct organismal differentiation and development.

5.2 Developmental syndromes associated with cohesin mutations

Mutations in the cohesin pathway cause multispectrum developmental disorders collectively known as “cohesinopathies”. The most characterized ones are Cornelia de Lange syndrome (CdLS) and Roberts syndrome (RBS). They entail clinical phenotypes that include intellectual deficiency, growth retardation, craniofacial dysmorphism, limb defects and cardiac malformations (Piché *et al.* 2019). CdLS affects 1 individual in every 10,000 live births. Most cases of CdLS are caused by inactivating mutations in the gene encoding NIPBL, causing haploinsufficiency, but mutations in SMC1A, SMC3, RAD21 and HDAC8 are also found to a lesser extent. Cells from CdLS patients or mouse models of CdLS do not exhibit obvious cohesion defects but have altered transcription (Kawauchi *et al.* 2009; Remeseiro *et al.* 2013a). RBS is rarer and caused by homozygous mutations in the gene encoding Esco2, a cohesin acetyltransferase. Unlike for CdLS, RBS patients exhibit premature centromere separation, aneuploidy and other defects related to cohesion loss (Whelan *et al.* 2012). Despite arising from mutations in a common cohesin pathway and having overlapping phenotypes, the underlying mechanisms for both syndromes are currently believed to be distinct.

More recently, a new genetic disorder with strong resemblance to CdLS was discovered: CHOPS syndrome (C for cognitive impairment and coarse facies, H for heart defects, O for obesity, P for pulmonary involvement and S for short stature and skeletal dysplasia). This syndrome is caused by gain-of-function mutations in the gene encoding AFF4, a member of the super elongation complex (SEC) that mobilizes paused RNA polymerase II. The mutations cause aberrant genomic distribution of both AFF4 and cohesin around

active genes, which was also observed in CdLS, in both cases leading to similar transcriptional deregulation (Izumi *et al.* 2015). The similarities could point to a common mechanism of pathogenesis in both syndromes related with the role of cohesin in gene regulation.

In recent years, clinical exome sequencing (CES) has revealed de novo germline mutations in *STAG1* and *STAG2*, genes encoding SA1 and SA2, respectively. The clinical features of the patients partially overlap with those of CdLS and other cohesinopathies, thus further expanding the genetic heterogeneity underlying these syndromes. For *STAG1*, 20 cases have been described with heterozygous mutations or microdeletions of the gene (Lehalle *et al.* 2017; Yuan *et al.* 2019). For *STAG2*, a total of 22 patients with mutations have been identified. The majority of them are female, but the fact that it is an X-linked gene has to be considered. Mutations in females are carried in heterozygosis and can be either missense or truncating (Aoi *et al.* 2019; Kruszka *et al.* 2019; Mullegama *et al.* 2017; Yuan *et al.* 2019). However, male patients with *STAG2* mutations have also been identified. These are carried in hemizygosis and are exclusively missense (Soardi *et al.* 2017; Yuan *et al.* 2019). A single case of a truncating mutation was found in a foetus that was lost at mid-pregnancy (Aoi *et al.* 2019). These findings are consistent with the prediction that truncating variants of X-linked genes impose more severe pathogenic effects on males than females. Data indeed indicate that truncations in *STAG2* are indeed less tolerated in males.

In addition to mutations, copy number variations (CNVs) have also been described for *STAG2* through genome-wide array comparative genome hybridization (CGH) or X-chromosome exome sequencing (Bonnet *et al.* 2009; Di Benedetto *et al.* 2014; Kumar *et al.* 2015; Leroy *et al.* 2016; Philippe *et al.* 2013; Yingjun *et al.* 2015). The CNVs consist of (micro)duplications or triplications of Xq25 in which the shortest region of overlap contains *STAG2*. These cases show that increased dosage can also lead to a cohesinopathy phenotype.

Overall, patients with mutations in *STAG1* or *STAG2* share the core clinical findings of cohesinopathies, including intellectual disability, developmental delay and dysmorphism, but tend towards the mild end of the spectrum. Their phenotypes are highly heterogeneous and the clinical severity likely depends on numerous factors, such as the type and position of the mutation, and the X chromosome inactivation (XCi) process in females in the case of *STAG2* (Mullegama *et al.* 2019).

There is still a limited number of cases but *STAG1* and *STAG2* should be considered as new cohesinopathy genes. Identification of novel patients can further expand the phenotype and provide a platform for functional studies.

5.3 Cohesin mutations in cancer

Cohesin was first associated with cancer when mutations in genes encoding cohesin subunits and its loader were identified by targeted sequencing in colorectal cancer (Barber *et al.* 2008). A few years later, loss of function mutations in *STAG2* were found in glioblastoma, Ewing sarcoma and melanoma samples (Solomon *et al.* 2011). This study showed that correction of mutant alleles in glioblastoma cell lines restored chromosome cohesion defects without altering the transcriptional profile, hinting to mis-segregation as the main pathogenic consequence of cohesin dysfunction. More recently, large projects of cancer genome sequencing have identified recurrent somatic mutations in genes encoding cohesin subunits or its regulators across many cancer types, pointing to the cohesin network as one of the most frequently mutated in cancer (Kandoth *et al.* 2014; Lawrence *et al.* 2014; Leiserson *et al.* 2015; Martincorena *et al.* 2017) (Fig. I12). Contrary to initial theories, more recent functional studies indicate that most tumors or cancer cell lines analyzed do not show a clear correlation between cohesin mutational status and aneuploidy (Balbas-Martinez *et al.* 2013; Kim *et al.* 2016; Welch *et al.* 2012). This suggests that cohesin dysfunction contributes to cancer pathogenesis through alternative mechanisms, possibly transcriptional deregulation.

Overall, *STAG2* harbored a significantly higher proportion of inactivating mutations than other genes in the cohesin network and was appointed as one of twelve genes significantly mutated in at least four tumor types, along with well-established cancer genes such as TP53, PTEN and KRAS (Lawrence *et al.* 2014). *STAG2* mutations were particularly prevalent in bladder cancer, myeloid malignancies and Ewing sarcoma but occur in a wide spectrum of cancer types.

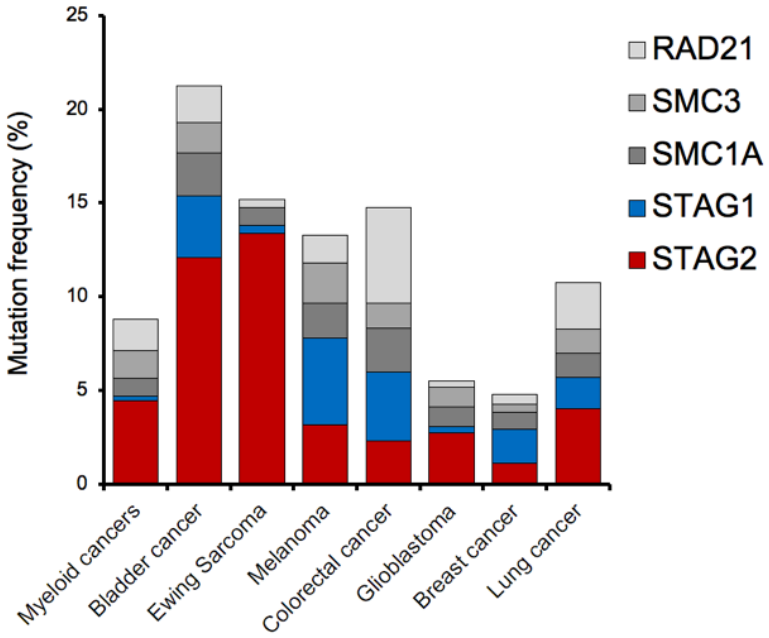


Figure I12. Frequency of mutation in genes encoding cohesin core subunits in cancer. The cohesin pathway, including core cohesin subunits as well as cohesin regulatory factors, is frequently mutated in multiple cancer types. Data obtained from cBioportal.

Although the prognostic impact of cohesin mutations has long been controversial (Kihara *et al.* 2014; Thol *et al.* 2014; Thota *et al.* 2014), retrospective analysis found that cohesin mutations trend towards a worse prognosis (Haferlach *et al.* 2014; Thol *et al.* 2014; Thota *et al.* 2014). Clonal analysis showed that cohesin mutations are present in the major tumor population, indicating they are early events in tumorigenesis (Kon *et al.* 2013; Thota *et al.* 2014), and discrimination by variant allele frequency suggests that mutations are in the dominant clone but not in the founder clone (Jan *et al.* 2012; Thol *et al.* 2014). These findings indicate that cohesin mutations could represent a transformative event towards more aggressive disease or that they provide a specific genetic context for the acquisition of additional mutations that promote tumorigenesis.

Many efforts have been made towards understanding the role of cohesin mutations in myeloid cancers. Sequencing studies have revealed a prevalence close to 10% of cohesin mutations in acute myeloid leukemia (AML) and other myeloid malignancies. Functional studies have shed light on the contribution of cohesin mutations to tumorigenesis. When *SMC1A* and *RAD21* AML-derived mutations were introduced in hematopoietic stem and progenitor cells (HSPCs), there were no changes in proliferation but differentiation was impaired. Increased accessibility was observed in regions

enriched for DNA-binding motifs of ERG, GATA2 and RUNX1, transcription factors involved in maintenance of the HSPC stemness programs (Mazumdar *et al.* 2015).

Two additional studies have explored the consequences of cohesin knock-down in hematopoiesis, *in vivo* using shRNAs against SMC1A, Rad21 and SA2 and a conditional knock out of SMC3. SMC3 haploinsufficiency resulted in increased self-renewal of HSPCs and reduced expression of lineage commitment genes (Mullenders *et al.* 2015; Viny *et al.* 2015).

Despite differences in model systems, taken together these studies strongly suggest that decreased cohesin levels may promote tumorigenesis through delaying or skewing HSPC differentiation and enforcing stem-cell programs instead. They appear to do so through modulation of chromatin accessibility to transcription factors involved in stem-cell maintenance. Importantly, reduced cohesin dosage by itself did not result in AML, but it enhanced tumorigenesis when combined with other alterations. This is consistent with the idea that cohesin mutations cooperate with additional mutations to promote malignancy.

These findings, together with the lack of correlation with aneuploidy, provide mounting evidence to support a mechanism whereby cohesin mutations alter 3D genomic organization and gene regulation. In the case of SA2-deficient tumors, SA1 likely allows survival by ensuring sufficient cohesion, but in turn cannot compensate SA2-specific roles, leading to a loss in the fidelity of chromatin loops with important functional consequences.

Objectives
Objetivos

Objectives

The main goal of this Thesis was to further clarify the functional specificity of somatic cohesin variants: cohesin-SA1 and cohesin-SA2. To this end, we generated a novel knock out mouse model for SA2 to serve as a tool for *in vitro* and *in vivo* studies. More specifically, our aims were:

1. Assess effects of SA2 depletion on cohesion-related functions in mouse embryonic fibroblasts (MEFs).

2. Address the genome-wide distribution of cohesin-SA1 and cohesin-SA2 and their contribution to gene regulation in MEFs.

3. Evaluate the role of cohesin-SA2 in murine embryonic development.

Objetivos

El objetivo principal de esta Tesis era esclarecer en mayor detalle la especificidad funcional de las variantes somáticas de cohesina: cohesina-SA1 y cohesina-SA2. Para ello generamos un nuevo modelo mutante de ratón para SA2 como base para estudios *in vitro* e *in vivo*. Nuestros objetivos específcos eran:

1. Estudiar los efectos de la depleción de SA2 en funciones relacionadas con la cohesión en fibroblastos embrionarios de ratón (MEFs).

2. Caracterizar la distribución de la cohesina-SA1 y la cohesina-SA2 a lo largo del genoma y su contribución a la regulación génica.

3. Evaluar el papel de la cohesina-SA2 durante el desarrollo embrionario de ratón.

Materials and methods

Materials and methods

***Stag2 cKO* mouse model generation.** The targeting vector PG00032_A_D11-3 was obtained from EUCOMM and electroporated into G4 mouse embryonic stem cells. Clones were selected in G418 and screened by Southern blotting for homologous recombination. Positive clones were infected with adeno-FLP to remove the selection cassette and create the conditional allele and microinjected into C57BL/6BrdCrHsd-Tyr morulae (CNIO Transgenic Mice core unit). Germline transmitting chimeras were screened by PCR (using primers Stag2_F and Stag2_R1; see Table M1) and selected to generate the mice colonies. More details in Results section 1 and Fig. R1.

Mouse housing. Mice were housed in a pathogen-free animal facility following the animal care standards of the institution. All animal procedures were approved by local and regional ethics committees (Institutional Animal Care and Use Committee and Ethics Committee for Research and Animal Welfare, Instituto de Salud Carlos III) and performed according to the European Union guidelines.

Mice strains. Mice carrying the *Stag2 cKO* allele were crossed with mice carrying *Cre* transgenes under the control of different promoters: *Tg.hUBC-CreERT2* (Ruzankina *et al.* 2007) for MEF isolation and *in vitro* studies, and *Tg.CAG-Cre* (Belteki *et al.* 2005) 2005) and *Tg.Sox2-Cre* (Hayashi *et al.* 2002) for embryo analyses. All mice strains were maintained in a predominantly C57BL/6 background.

Timed matings. Male and female mice of reproductive age were placed together overnight. The next morning, females were checked for vaginal plugs to determine if mating had occurred. If the female was pregnant, time of plug detection was considered day E0.5 of embryonic development. Embryos were collected between E8.5 and E12-5, after palpation of females to confirm pregnant status.

MEF isolation, culture and immortalization. Primary MEFs were isolated from E12.5 embryos. Pregnant females were sacrificed in a CO_2 chamber, uterine horns were dissected and transferred to a sterile PBS solution. Embryos were extracted from the uterus in a laminar flow hood. Fetal liver was excised and the head was taken for subsequent genotyping (using Stag2_F+R and CreERT2_F+R; see table M1). The rest of the embryonic tissue was minced with a scalpel and treated with 0.25% trypsin-EDTA (Sigma-Aldrich) for 10 min at 37ºC. Cells were further disgregated by pipetting and

resuspended in 9 mL medium. Primary MEFs were routinely grown in DMEM (Lonza) supplemented with 20% FBS (Sigma-Aldrich) and 1% penicillin-streptomycin, in an incubator at 37ºC under 90% humidity and 5% CO_2. All experiments were performed in primary low passage MEFs except for FRAP, for which MEFs were immortalized using the SV40 large T antigen. Immortalized MEFs were grown in DMEM with 10% FBS and antibiotics.

Stag2 ablation in culture. *Stag2 cKO* MEFs were treated with 1 μM 4-hydroxy tamoxifen (4-OHT) for 4 days in asynchronous cultures (20% FBS) and for 3 days in serum-starved conditions (0.1% FBS). The efficiency of depletion was assessed by immunoblotting (antibodies in Table M2). The same cells cultured without 4-OHT served as control.

Immunoblotting. Whole cell extracts for immunoblot were prepared by lysing in Laemmli buffer at 10,000 cells/μl, sonicating and boiling for 5 min at 95ºC. SDS-polyacrylamide gels and immunoblotting were performed following standard protocols. Custom-made primary antibodies were used at 2 μg/mL for 1 hour/RT and commercial antibodies following the manufacturer instructions (Table M2). Horseradish peroxidase (HRP)-conjugated secondary antibodies (Amersham Biosciences) were used at 1:5000 dilution in 5% Milk for 1 hour/RT. ECL developing reagent (Amersham Biosciences) was used.

Proliferation assays. MEFs pretreated for 4 days with 4-OHT were seeded at 11,000 cells/cm^2 in multiwell plates (3 triplicate wells per timepoint). Over the following days, cells were collected and counted in a Neubauer hemocytometer.

Cell cycle analysis by FACS. MEFs grown at different conditions with or without 4-OHT were collected after a 30 min pulse with 30 μM BrdU (Sigma) and fixed overnight in 70% ethanol. DNA was denatured with 2N HCl for 20 min at RT. Cells were blocked with 1% BSA-0.05% Tween20 in PBS. Cells were incubated with a FITC-conjugated anti-BrdU antibody (Table M2) for 1h at 37ºC and DNA was stained overnight with 50 μg/ml propidium iodide (Sigma) with 10μg/mL RNase A (Qiagen). Flow cytometry was performed in a FACS Canto II cytometer (BD) and profiles were analyzed using FlowJo 10.0.8 software.

Cohesion and segregation analyses. To enrich MEFs in mitotic cells, they were serum-starved for 3 days in media with 0.1% FBS in the presence or absence of 4-OHT and

released into media with 20% FBS for 36h. For chromosome spreads, 0.1 µg/ml colcemid was added to the medium 3-4 h before harvesting. Cells were swollen in 0.03M sodium citrate, fixed in methanol:acetic acid 3:1 and dropped onto slides. For anaphase analysis, cells were seeded onto coverslips at the time of release from G0 arrest. In both cases, cells were stained with 1 µg/ml DAPI, mounted with Vectashield and imaged using a Leica DM6000 microscope with LAS AF software.

Chromatin Immunoprecipitation (ChIP). ChIP was performed in asynchronously growing MEFs as described (Remeseiro *et al.* 2012b). For each condition, two clones from different embryos were processed independently. Cells were cross-linked with 1% formaldehyde for 15 min at RT and quenched with 0.125M glycine. After two washes in PBS with 1 µM PMSF and protease inhibitors, cells were scraped, pelleted and lysed in lysis buffer (1% SDS, 10mM EDTA and 50mM Tris-HCl) at 20,000 cells/µL. Chromatin was sonicated in a Covaris system (shearing time 30 min, 20% duty cycle, intensity 6, 200 cycles per burst, 30 sec per cycle). 25 µg of antibody was used to immunoprecipitate from 50 µg of sheared chromatin. Around 5 µg of immunoprecipitated chromatin per sample was used for library preparation. Adaptor-ligated libraries were generated by limited-cycle PCR with Illumina PE primers and were sequenced on an Illumina Genome Analyzer IIx platform.

ChIP-sequencing analysis. Alignment of sequences to the reference mouse genome (mm9, February 2009) was performed using 'Bowtie2′ (version 2.3.3.1) under default settings (Langmead & Salzberg 2012). Duplicates were removed using Picardtools (version 2.13.2) and peak calling was carried out using MACS2 (version 2.1.1.20160309) after setting the q-value (FDR) to 0.05 and using the '–extsize' argument with the values obtained in the 'macs2 predictd' step (Zhang *et al.* 2008b). Mean read-density profiles and read-density heatmaps for different chromatin-binding proteins were generated with deepTools 2.5.4 (Ramírez *et al.* 2016). In addition to data generated in this study and in our group, we used publicly available datasets from Tedeschi *et al.* 2013 (GSE41603), Busslinger *et al.* 2017 (GSE76303) and ENCODE data for chromatin states (GSM1000139, GSM769028, GSM656318). Chromatin states were defined by exclusive presence of the following histone marks: H3K27ac and H3K4me1 for strong enhancers, H3K4me1 for weak enhancers, H3K27ac and H3K4me3 for strong promoters, H3K4me3 for weak promoters, CTCF for insulators. "Rest" chromatin state contained none of the above. Motif enrichment analysis was performed with HOMER, using default parameters that screen the 200bp region surrounding the ChIP-seq peak summits.

Generation of cell lines with GFP-tagged cohesin subunits. One parental clone of immortalized MEFs was used to generate Rad21-, SA1- and SA2-GFP cell lines by CRISPR-Cas9 as described (Ladurner *et al.* 2016). Donor plasmids containing the C-terminus of the targeted genes with in-frame GFP were created by Gibson Assembly. sgRNA sequences were designed using crispr.mit.edu (Table M3) and cloned in pX335 plasmids, that also encode Cas9n-D10A. Plasmids were introduced into MEFs by electroporation with a Neon Transfection System (Thermofisher) applying 2 pulses of 20 ms at 1400V. Positive cells were selected through an Influx Cell Sorter (BD) based on the GFP signal over control cells and the resulting polyclonal population was characterized by immunoblot and immunoprecipitation. More information in Results section 3.4 and 3.5.

Antibody generation. New custom-made antibodies against the N-terminal end of SA1 and SA2 were generated as part of this study, by using recombinant proteins with His-MBP tag expressed in *E. coli*. Rat monoclonal antibodies were raised against residues 1-223 of SA1 (CNIO Monoclonal Antibodies Unit). Rabbit polyclonal sera were obtained from rabbits injected with a fragment containing aminoacids 1-220 of SA2 and affinity purified.

Biochemical fractionation and salt extraction. We followed the protocol from Méndez & Stillman 2000. Cells were resuspended at $2 \cdot 10^7$ cells/mL in buffer A (10 mM HEPES pH 7.9, 10 mM KCl, 1.5 mM $MgCl_2$, 0.34 M sucrose, 10% glycerol, 1 mM DTT, 1 mM $NaVO_4$, 0.5 mM NaF, 5 mM β-glycerophosphate, 0.1 mM PMSF), and incubated on ice for 5 min in the presence of 0.1% Triton X-100. Low-speed centrifugation (4 min/600 g/4°C) allowed the separation of the cytosolic fraction (supernatant) and nuclei (pellet). Nuclei were washed and subjected to hypotonic lysis in buffer B (3 mM EDTA, 0.2 mM EGTA, 1 mM DTT, 1 mM $NaVO_4$, 0.5 mM NaF, 5 mM β-glycerophosphate, 0.1 mM PMSF) 30 min on ice. Nucleoplasmic and chromatin fractions were separated after centrifugation (4 min/600 g/4°C). Chromatin was resuspended in Laemmli Sample Buffer and sonicated twice for 15 seconds at 20% amplitude. For salt extraction experiments, chromatin fractions were either left untreated or treated with 0.5M NaCl in modified buffer A (10 mM HEPES pH 7.9, 1.5 mM $MgCl_2$, 0.34 M sucrose, 10% glycerol and supplemented as above) for 30 min on ice. Solubilized proteins were separated from insoluble chromatin by low-speed centrifugation (4 min/600 g/4°C) and prepared for immunoblotting.

Immunoprecipitation, Asynchronously growing cells were lysed on ice for 30 min in lysis buffer [0.5% NP-40 in TBS supplemented with 0.5mM DTT, 0.1mM PMSF and 1X complete protease inhibitor cocktail (Roche)] and sonicated. Then NaCl was added to 0.3M and the extract rotated for 30 min at 4ºC. Salt concentration was the lowered to 0.1M NaCl by dilution and glycerol added to 10% final concentration. Extracts were incubated with the specific antibodies for 2h at 4ºC and rotated with 1/10 volume of protein A agarose beads for 1h at 4ºC. The beads were washed 6 times with 20 volume of lysis buffer and eluted in SDS-DTT gel loading buffer for 5min at 95ºC.

Immunofluorescence in cultured cells. MEFs grown on coverslips were fixed with 4% PFA for 20 min and permeabilized in 0.25% Triton X-100 in PBS for 5 min on ice. Cells were blocked with 3% BSA, 0.05% Tween 20 in PBS for 30 min. Primary and secondary antibodies were diluted in blocking solution and incubated for 1h each. DNA was counterstained with 1 mg/ml DAPI. A Leica DM6000 microscope was used to obtain grayscale images, which were later analyzed using FIJI software.

iFRAP. Cells were seeded in 8-well chambered coverslips (Ibidi) at 40,000 cells/cm^2 48h prior to performing the experiment. The next day media was changed to 0.1% FBS for 24h. iFRAP was performed in a Leica TCS-SP5 (AOBS) confocal microscope from Germany Leica Microsystems using a 40x/1.2 NA HCX PL APO objective with immersion oil. Cells were kept in a climate chamber at 37ºC with 5% CO2 during the experiment. Image acquisition used the HCSA software in LAS AF 2.7. Cells were photobleached with an argon laser and the recovery was monitored by live-cell imaging, Pictures were taken immediately before and after photobleaching as well as every 30 seconds during recovery. Videos were analyzed using FIJI software and statistical analysis and non-linear regression with GraphPad Prism.

RNA extraction from MEFs for RNA-sequencing. We collected 3 paired clones of asynchronous growing MEFs per condition. RNA was extracted with the RNeasy kit (Qiagen). PolyA+RNA was purified with the Dynabeads mRNA purification kit (Invitrogen), randomly fragmented and converted to double-stranded cDNA and further processed as in Illumina's TruSeq RNA Sample Preparation Guide. Adapter-ligated libraries were made by limited-cycle PCR with Illumina PE primers and sequenced on HiSeq2000 or HiSeq2500 platform.

RNA-sequencing analysis. Fastq files with single-end sequenced reads were quality-checked with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ by S.

Andrews) and aligned to the mouse genome (mm9) with Nextpresso (Graña *et al.* 2018) executing TopHat-2.0.0 using Bowtie 0.12.7 and Samtools 0.1.16 allowing two mismatches and five multi-hits. Transcript assembly, estimation of their abundances and differential expression were calculated with Cufflinks 1.3.0 using the mouse genome annotation dataset NCBIM37/mm9 from Ensembl. To account for multiple-hypothesis testing, the estimated significance level (P value) was adjusted using Benjamini–Hochberg FDR correction. We consider changes with FDR<0.05 and FPKM>2 as significant. GSEAPreranked was used to perform a gene set enrichment analysis (using also Nextpresso).

Embryo extraction and processing for histology and immunofluorescence. Whole mount embryos were dissected in PBS at RT and imaged using a LEICA MZ10F microscope and LAS 3.8 software. DNA from the embryos or their yolk sac was extracted to genotype (using Stag2_F+R1, Stag2_F+R2, Sry_F+R, see Table M1). Embryos were fixed in a 10% formalin solution at pH 7 (Sigma HT501128-4L) overnight at 4ºC, dehydrated in an ethanol series and stored in ethanol 70% at 4ºC until further processing. Embryos were embedded in paraffin and sectioned at 5µm, longitudinally for embryos at E8.5 stage and transversely for embryos at E9.5 and E10.5. H-E was performed according to standard procedures and imaged with a Nikon Eclipse 90i microscope and NIS Elements D3.2 imaging software. Histology was evaluated in 2-4 embryos per genotype based on consecutive sections [WT1 (n=3), KO mild (n=4), WT2 (n=3) and KO severe (n=2) at E9.5; WT (n=3) and KO (n=3) at E10.5. For neural tube and aortas, section at the height of the heart chambers were taken.

Immunofluorescence staining of embryo sections. Co-immunostaining for H3P-TUNEL-ISL1 (Table M1) was performed on 4 embryos per genotype at E9.5 (WT1, KO and WT2). For TUNEL, the Terminal Transferase recombinant kit (Roche 03 333 574 001) and biotin-16-dUTP (Roche 11 093 070 910) were used. Sections were imaged with a Nikon A1R confocal microscope and NIS Elements 4.30 software. For analysis, 2-3 non-consecutive sections were analyzed per embryo. H3P signal was quantified with a custom-made Image J macro, taking into account both late G2 and M-phase signals. Single ISL1 staining was performed on 4 embryos per genotype at E8.5 (WT and KO) and imaged as above. 4 sections were analyzed per embryo. Statistical significance was determined by Kruskal-Wallis test and Dunn's Multiple Comparison post-test using GraphPad Prism 5.03.

Embryo dissection, RNA extraction and RNA-sequencing. Whole-mount embryos were dissected in cold PBS. We used E9.5 WT and KO embryos with 21-23 pairs of somites. Heart tissue (whole heart along with surrounding SHF regions) and neural tube tissue (a section of heart-proximal neural tube) were dissected under a LEICA MZ10F microscope. Samples were immediately snap-frozen and stored at -80ºC. Per genotype and region, 3 replicates from 3 embryos each were pooled using TRI reagent (Sigma T9424) and homogenized by multiple passages through a syringe and needles (25-30G). Chloroform and phase lock tubes (QuantaBio 2302830) were used for phase separation and a subsequent precipitation with ethanol was performed at -20 ºC. RNA samples were analyzed using a Bioanalyzer 2100 (Agilent) and the RNA 6000 Pico kit. Libraries were prepared using the QuantSeq 3' mRNA-seq Library Prep Kit FWD (Lexogen) and sequenced (10 million single reads per sample) on an Illumina HiSeq 2500 platform.

RNA-sequencing analysis. For alignment and gene counting, we applied the Lexogen QuantSeq 2.2.3 pipeline provided by BlueBee, designed for use with the libraries described above. We removed one of the WT heart replicates due to initial inferior RNA integrity and a failure to cluster with the rest of the WT heart samples. The differential expression analyses have been performed with DeSeq2, excluding genes with no reads in any of the samples. Results were filtered by p-value<0.05 and FDR<0.05. In the heatmaps, color intensities correspond to the relative expression levels for each gene among conditions, normalized using the mean and standard deviation. Gene Ontology (GO) Enrichment Analysis was performed using Panther (Mi *et al.* 2013), using only gene sets with less than 1,000 genes to exclude the most general GO terms. GSEAPreranked was used to perform a gene set enrichment analysis (Subramanian *et al.* 2007) using the normalized counts provided by DeSeq2 using the default parameters.

Table M1. Primers for genotyping mice and embryos.

Primer	Sequence 5'-3'
Stag2_F	TGGTGCTTGGGATCAGATTT
Stag2_R1	TCCCTCATCAAAGTCGAAAA
Stag2_R2	AACAGCCTGAGCAAAGAATCC
Sry_F	TGGGACTGGTGACAATTGTC
Sry_R	GAGTACAGGTGTGCAGCTCT
CreERT2_F	TGAAGCTCCGGTTTTGAACT
CreERT2_R	GGTTCTTGCGAACCTCATCAC

Table M2. Antibodies used in this study.

Antibody	Reference	Use
SA1 rat monoclonal (Nt)	This study, Kojic *et al.* 2018	Undiluted supernatant for immunoblotting (IB)
SA1 rabbit polyclonal	Remeseiro *et al.* 2012b	2 µg/ml for immunofluorescence (IF)
SA2 mouse monoclonal	sc-81852 (SCBT)	1:100 for IF
SA2 rabbit polyclonal (Nt)	This study	2 µg/ml for IB
Rad21 rabbit polyclonal	Carretero *et al.* 2013	2 µg/ml for IB
MEK2 mouse monoclonal	M24520 (BD)	1:2000 for IB
BrdU-FITC	556028 (BD)	1:50 for flow cytometry
SMC1	Remeseiro *et al.* 2012b	for IP and ChIP
CTCF	07-729 (Millipore)	for ChIP
H3	1791 (Abcam)	1:10000 for IB
ISL1 mouse monoclonal	39.4D5 (DSHB Hybridoma Bank)	1:100 for IF
H3P rabbit polyclonal	06-570 (Millipore)	1:200 for IF

Table M3. sgRNA sequences used for CRISPR.

Primer	Sequence 5'-3'
SA1-sg1	CACCGTCTTCAGACTTCAGAACAT
SA1-sg2	CACCGCTCTATTATTTAGAGCTAG
SA2-sg1	CACCGCCCAAGAACCTAATGAGAG
SA2-sg2	CACCGCAGTACACATTTAAATCTG
Rad21-sg1	CACCGGTGACATCATTGCAACCCC
Rad21-sg2	CACCGTCTAGCTCCTCAGATAATA

Results

Results

1. Generation of a conditional knock out mouse model for SA2

In order to explore the role of cohesin-SA2 on multiple levels, we decided to generate a conditional knock out (cKO) mouse model for *Stag2*. For this purpose, we purchased a targeting vector from the European Conditional Mouse Mutagenesis program (EUCOMM) that allows the generation of a knock out first allele with conditional potential. The vector provides homology arms of 4.1 and 5.2 kb that promote recombination into the *Stag2* locus, located on chromosome X, where it targets exon 7 out of 38. In the sequence corresponding to intron 6 it contains an FRT-flanked cassette with a splicing acceptor site (SA), a neomycin resistance selection gene (*neo*) and a polyadenylation sequence (pA). In addition, the vector contains exon 7 flanked by loxP sites (Fig. R1A). When the FRT-flanked cassette is present, transcription is terminated at the pA, generating a null allele ($Stag2^{frt}$). This cassette can be excised by action of the recombinase Flippase (Flp), which recognizes and recombines the FRT (Flp Recognition Target) sequences. This leaves only the loxP sites flanking the targeted exon, thus creating a conditional allele ($Stag2^{lox}$). These sites can be recognized and recombined by another recombinase, Cre. When this excision occurs, a null allele is produced ($Stag2^{\Delta}$; Fig. R1B).

The targeting vector was introduced in mouse embryonic stem cells (mESCs) by electroporation and clones were screened by Southern blotting using two probes (Fig. R1C). Correctly targeted clones were infected with adeno-Flp to eliminate the FRT cassette and convert the allele to its conditional configuration. Next, selected clones of ES cells were microinjected into blastocysts and implanted into pseudopregnant mothers. Chimeras with germline transmission were selected to establish a mouse colony and genotypes were routinely confirmed by PCR analyses of genomic DNA (Fig. R1D). This mouse model served us to obtain *Stag2* cKO mouse embryo fibroblasts (MEFs) to assess cohesin-SA2 function in cultured cells and to address its relevance for murine embryogenesis.

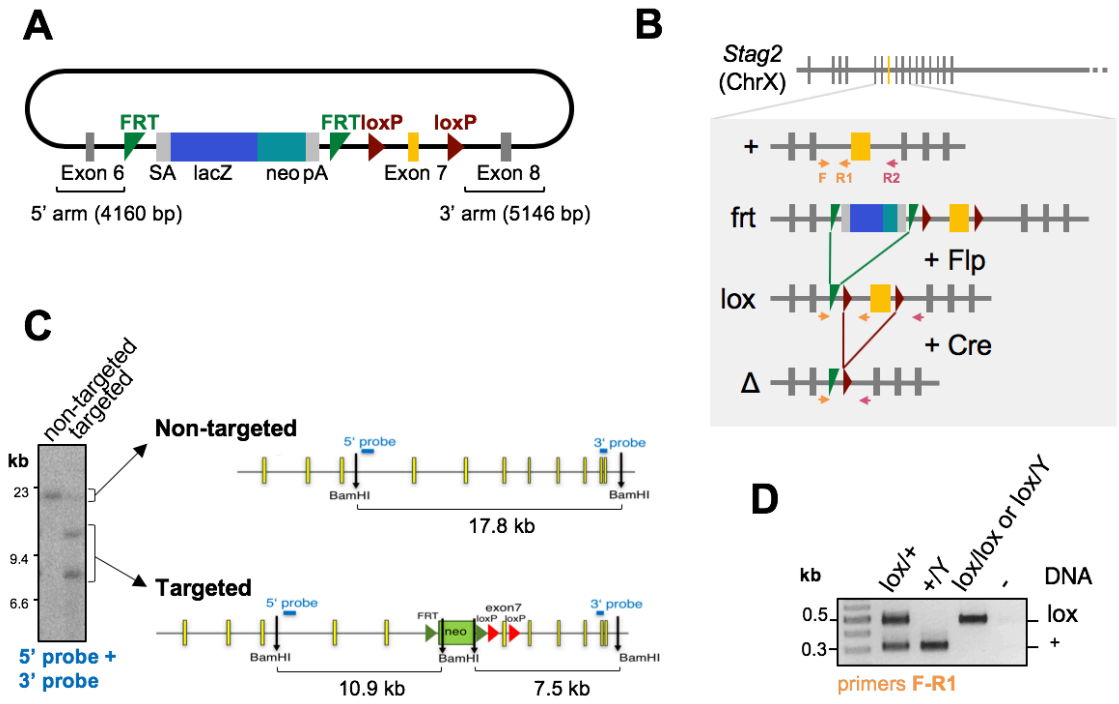


Figure R1. Generation of a novel knock-out mouse model for *Stag2* by gene targeting.

A. Map of the targeting vector obtained from EUCOMM to target the murine *Stag2* gene.

B. Scheme representing the targeting of the *Stag2* allele and its conversion to different configurations in the presence of flippase (Flp) or Cre recombinases: + (wild type), *frt* (knock out first), *lox* (conditional) and Δ (induced knock out). Primers used for genotyping are indicated.

C. Southern blot analysis (left) and strategy (right) for selection of correctly targeted mESC clones.

D. Genotyping PCR analyses to distinguish between the *Stag2* conditional (*lox*) and wild type (+) alleles in mice, using F and R1 primers.

2. Specific roles of cohesin-SA2 in cell proliferation

2.1 Generation of conditional knock out MEFs for SA2

To assess basic cohesin functions we chose an *in vitro* approach using cKO MEFs obtained from the *Stag2* mouse model described above. Conditional *Stag2*$^{lox/lox}$ females were mated with males carrying *hUBC-CreERT2* (Fig. R2A). This transgene ubiquitously expresses a fusion protein of Cre recombinase and a mutant form of the estrogen receptor that, in presence of tamoxifen, is activated by translocation to the nucleus (Ruzankina *et al.* 2007). We extracted embryos at embryonic stage E12.5 and obtained MEF cultures by mechanical and chemical disgregation (Fig. R2A). After genotyping *Stag2* and *CreERT2*, only male cKO MEFs (*Stag2*$^{lox/Y}$*; hUBC-CreERT2*$^{+/T}$) were kept. In these MEFs, SA2 depletion can be induced by addition of 4-hydroxy-tamoxifen (4-OHT) to the culture medium for 3-4 days.

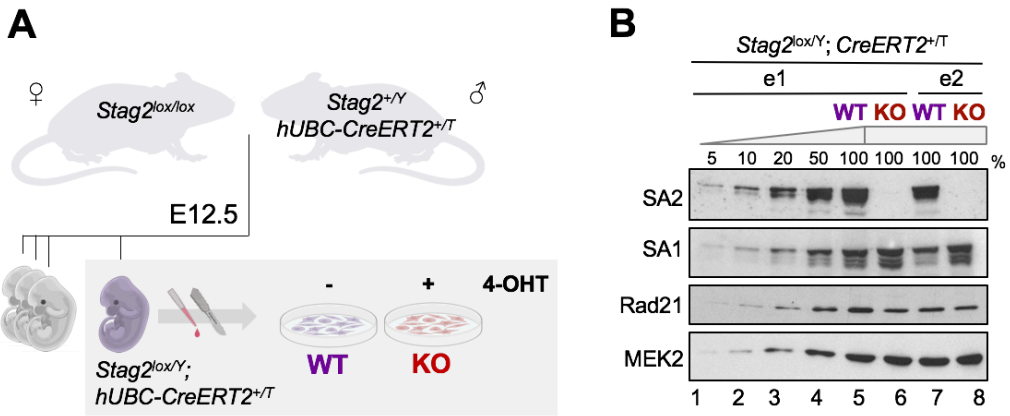


Figure R2. SA2 cKO MEFs as model for *in vitro* studies of cohesin variants.

A. Scheme for the establishment of cKO MEF cultures. *Stag2* conditional females were mated with males carrying one copy of the *hUBC-CreERT2* transgene and embryos at stage E12.5 were extracted to prepare MEFs. Male embryos carrying *CreERT2* were selected, in which SA2 depletion could be induced by addition of 4-OHT to the culture medium.

B. Immunoblot analysis of whole cell extracts from cKO MEFs derived from two different embryos (e1 and e2) untreated or treated with 4-OHT to deplete SA2 (WT and KO) for 4 days in culture. Decreasing amounts of WT MEF extract (shown as % of maximal) were loaded to estimate the remaining amount of SA2. MEK2 was used as a loading control.

SA2 depletion was efficient upon addition of 4-OHT for 4 days, and SA2 protein levels in treated MEFs (KO; lanes 6 and 8 in Fig. R2B) typically dropped below 5% of the amount present in untreated MEFs (WT; lanes 5 and 7 in Fig. R2B), as seen by immunoblotting in whole cell extracts. In most cases, we observed a compensatory upregulation of SA1 to variable extents. The efficient and reproducible depletion of SA2 made these cells a useful tool for further *in vitro* studies.

2.2 Slower proliferation and minor defects in S phase progression in MEFs depleted of SA2

Our first observation in this system was that SA2-null MEFs (KO hereafter) were viable, as cells kept growing in the almost complete absence of SA2. However, they proliferated at slower rates than their wild type (WT) counterparts (Fig. R3A). For this proliferation assay cells were seeded at low confluence and in these conditions, KO MEFs were particularly sensitive.

In order to test whether the reduced proliferation arises from problems in cell cycle progression we performed flow cytometry analysis. We pulse-labeled cells for 30 min with bromodeoxyuridine (BrdU), a synthetic thymidine analog which can be used together with the DNA stain to better distinguish cells in S phase (Fig. R3B). We did not find any significant difference in the cell cycle profiles of asynchronously growing WT and KO cells (Fig. R3C), suggesting they were not arrested at any given phase of the cell cycle. In all clones tested, a small but consistent reduction of BrdU-incorporating cells was observed among KO MEFs, but did not reach statistical significance (p=0.11). These data rather indicate that SA2-null MEFs cycle slower but mechanisms contributing to this phenotype remained unclear.

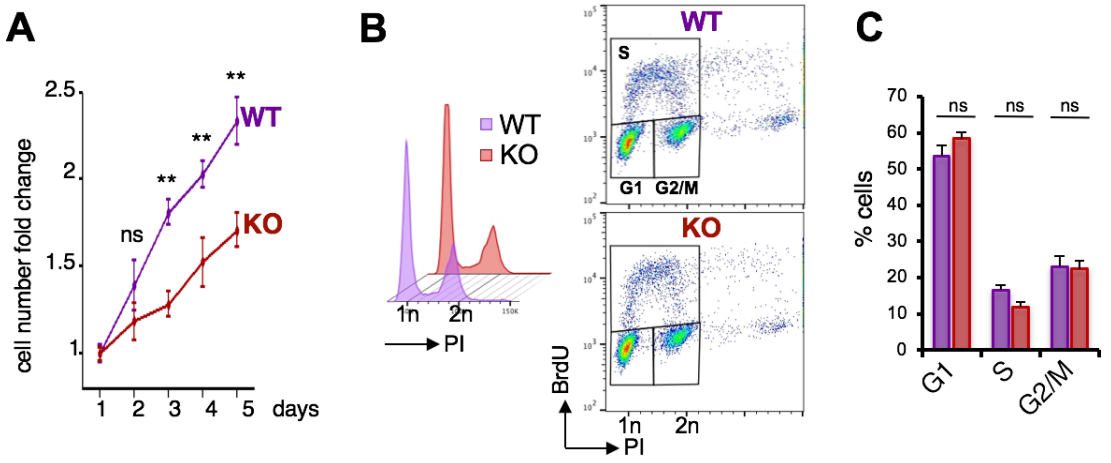


Figure R3. Slower proliferation of SA2KO MEFs and normal cell cycle distribution.

A. Growth curves of WT and KO MEFs representing the average fold increase in cell number relative to the number of cells seeded on day 1. Data from MEFs from 2 embryos (n=2), each analyzed in triplicates (mean ± SEM). Mann-Whitney test; ** P<0.01, ns P≥0.05.

B. Representative FACS profiles of asynchronously growing WT and KO MEFs. Propidium iodide (PI) histogram (left) and BrdU incorporation profiles with gated G1, S and G2/M populations (right).

C. Quantification (mean ± SEM) of cell cycle phases as gated on the BrdU incorporation profiles in panel B. n=4. Mann-Whitney test; ns P≥0.05.

A recent report has shown that human RPE cells lacking SA2 arrest in late S phase, supporting a role for cohesin-SA2 in replication fork progression (Mondal *et al.* 2019). The arrest becomes very pronounced over time, so we wondered whether treating MEFs with 4-OHT for longer times would result in a stronger phenotype. We compared the cell cycle profiles of MEFs treated for 4 and 8 days with 4-OHT (Fig. R4A). Cells were efficiently depleted of SA2 (Fig. R4B) but the fraction of BrdU-incorporating cells was reduced to a similar extent in SA2KO MEFs after 4 and 8 days of treatment (p=0.1; Fig. R4C and R4D).

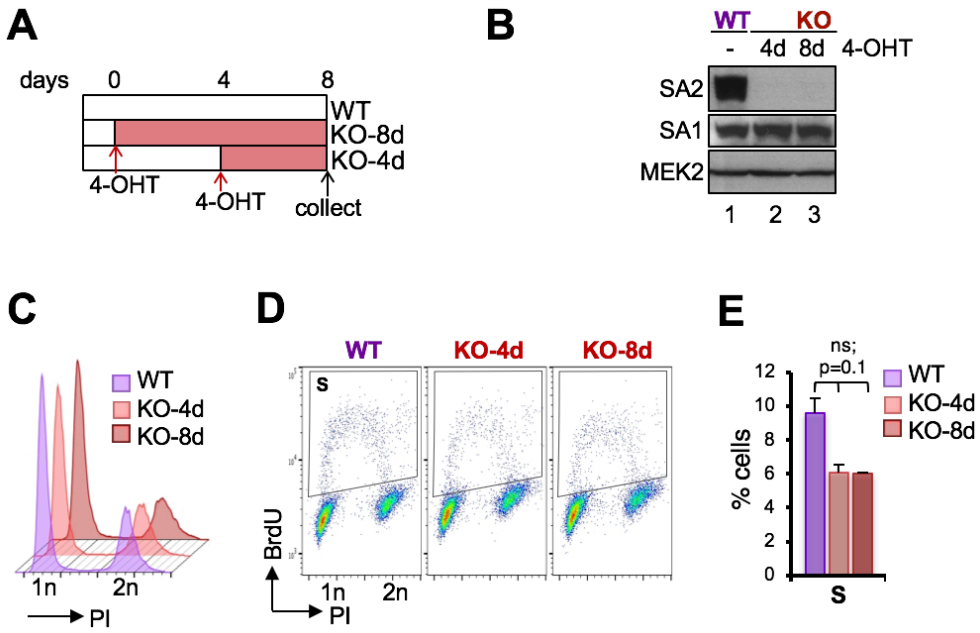


Figure R4. No obvious cell cycle changes upon longer treatment with 4-OHT.

A. Experimental scheme.
B. Immunoblot of whole cell extracts of WT and KO MEFs that have been treated for 4 or 8 days with 4-OHT. MEK2 is used as loading control.
C. Representative PI histograms by FACS.
D. Representative BrdU incorporation profiles.
E. Quantification (mean ± SEM) of cells in S phase, as gated in panel D. n=3. Mann-Whitney test; ns P≥0.05.

We also examined whether cells severely depleted of cohesin-SA2 had any problem in cell cycle re-entry and S phase initiation and progression after being released from a quiescent state. Serum starved *Stag2* cKO MEFs were cultured in the presence or absence of 4-OHT for 3 days, released in serum-rich media and monitored by flow cytometry after a BrdU pulse (Fig. R5A). MEFs with or without SA2 initiated S phase roughly at the same time (Fig. R5B and R5C), but the fraction of replicating cells was slightly smaller in the KO cells, especially at times 12 and 36 after release (p=0.07 and p=0.1, respectively).

Thus, SA2KO MEFs presented some minor defects in S phase progression that could contribute to their slower proliferation. We cannot discard that these defects might become more evident if S phase is challenged.

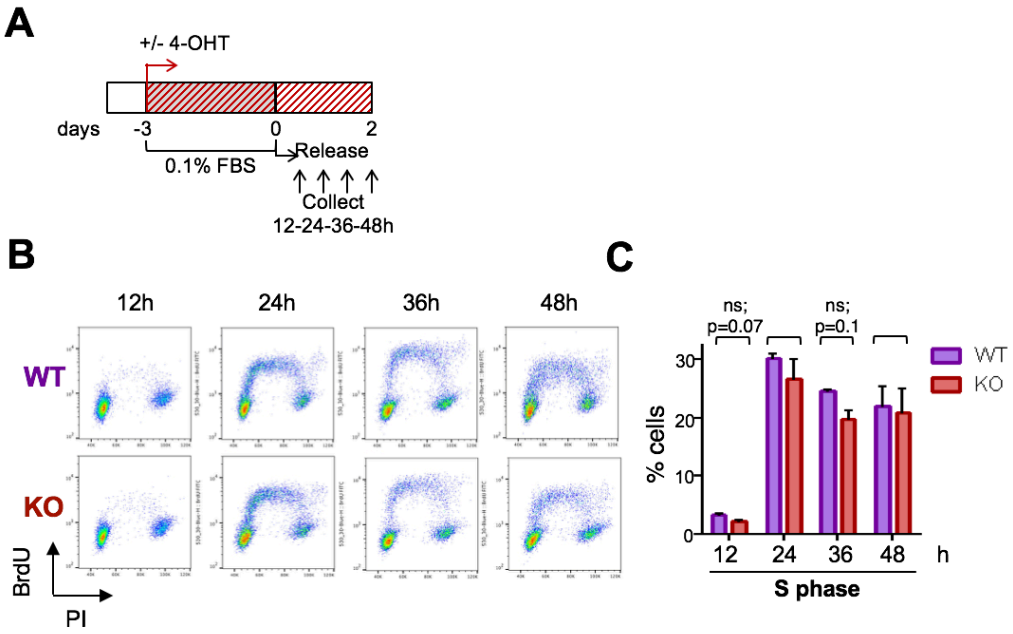


Figure R5. Cell cycle reentry after G0 arrest in SA2KO MEFs.

A. Experimental scheme.
B. Representative BrdU incorporation profiles from FACS analysis at 12-24-36-48h after release from G0.
C. Quantification (mean ± SEM) of BrdU-positive population (S phase). n=3. Mann-Whitney test; ns P≥0.05.

2.3 Cohesion defects in MEFs depleted of SA2

Cohesin was discovered for its role in cohesion between sister chromatids. This cohesion is established during S phase and lasts until mitosis, when two different waves of cohesin dissociation from chromatin promote proper chromosome segregation. As mentioned in the introduction, our lab and others have reported a more prominent role for cohesin-SA1 and cohesin-SA2 in telomere and centromere cohesion, respectively, in different cell types. Therefore, we explored the effect of SA2 depletion on cohesion in our particular cellular model, focusing on centromere cohesion.

Stag2 cKO primary MEFs were serum-starved for 3 days in presence or absence of 4-OHT and subsequently released back into serum-rich medium for 36 hours to give cells enough time to reach mitosis (Fig. R6A). We examined centromere cohesion in metaphase spreads and distinguished two major phenotypes: (1) “loosened” cohesion, evidenced by increased distance between sister centromeres compared to normal

chromosomes in which the two sister centromeres are fused and (2) “lost” cohesion, characterized by complete unpairing of sister centromeres (Fig. R6B). We detected very few cases of the more dramatic phenotype among WT and KO MEFs (1.3% and 3% respectively, dark green in Fig. R6C). We did, however, observe a larger fraction of chromosomes per metaphase with loosened centromeres in KO MEFs compared to WT (26 % and 11% respectively, light green in Fig. R6C).

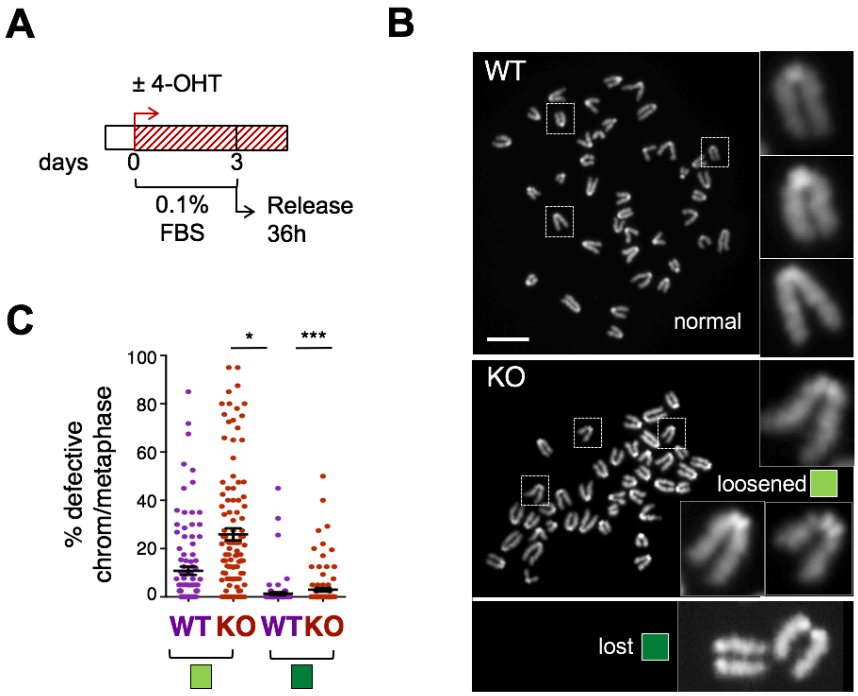


Figure R6. Centromere cohesion defects in MEFs depleted of SA2.

A. Representative metaphase spreads from WT and KO MEFs and examples of normal, separated and loosened centromeres.
B. Quantification of centromeric cohesion defects (mean ± SEM). Each dot represents a single metaphase. At least 100 metaphases from MEFs from 3 different embryos were inspected. Scale bar, 10 μm. Mann-Whitney test; *** P<0.001, * P<0.05.

These observations confirm a role for cohesin-SA2 in centromeric cohesion, as was previously described in other cell types. However, as the defects are very mild, it is clear that cohesin-SA1 and/or mechanisms contributing to cohesion other than cohesin provide sufficient cohesion at the centromere region of these cells in the absence of cohesin-SA2.

2.4 Chromosome segregation defects in MEFs depleted of SA2

To explore whether the mild centromere cohesion defects observed in cells lacking cohesin-SA2 have an impact on chromosome segregation, we looked at cells undergoing anaphase. As before, *Stag2* cKO MEFs were cultured in low serum in the

presence or absence of 4-OHT for 3 days and seeded onto coverslips at the time of release into serum-rich medium. After 36h, coverslips were collected, fixed and stained for DNA. Anaphase figures were inspected for lagging chromosomes and chromatin bridges (Fig. R7A). WT MEFs proved to be prone to segregation defects (17% of anaphases). Importantly, we found an almost 2-fold increase in defects in KO MEFs (29% of anaphases; Fig. R7B), even though data variability renders this difference statistically non-significant.

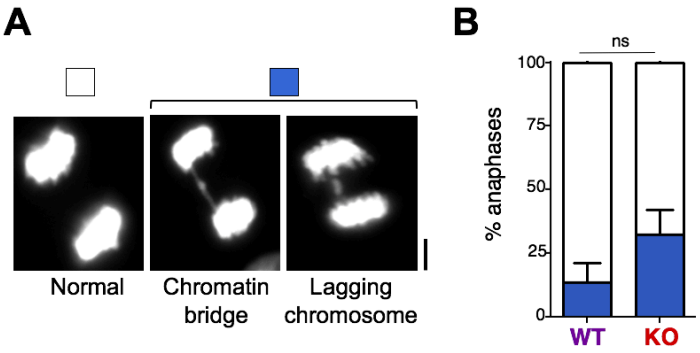


Figure R7. Increase in defective anaphases in SA2-deficient MEFs.

A. Images of anaphase cells, either normal or defective, found among WT and KO MEFs.
B. Quantification of anaphase defects (mean ± SEM), including lagging chromosomes and chromatin bridges. At least 100 anaphases from MEFs from 3 different embryos were inspected. Scale bar, 5 μm. Mann-Whitney test; ns P≥0.05.

Next, we analyzed chromosome number in metaphases. We prepared metaphase spreads from MEFs that had been cycling in the presence of 4-OHT for variable times: 36h after release from G0 (as in Fig. R6A), or 4 and 6 days in asynchronous conditions. In the first timepoint the number of chromosomes did not significantly differ from 40, which was to be expected as this cell population is only undergoing the first mitosis after release from G0 (Fig. R8A). In the case of KO MEFs treated for 4 days, we did find a significant deviation from the normal number of chromosomes (Fig. R8B), that was further increased after 6 days of treatment (Fig. R8C).

Overall, we concluded that SA2-null MEFs, despite displaying only mild cohesion defects, indeed present a lower fidelity of chromosome segregation. We do not rule out that these segregation defects have origins other than cohesion defects. In any case, the accumulation of aneuploidies may lead to inviable cells, which could also contribute to the observed reduction in proliferation.

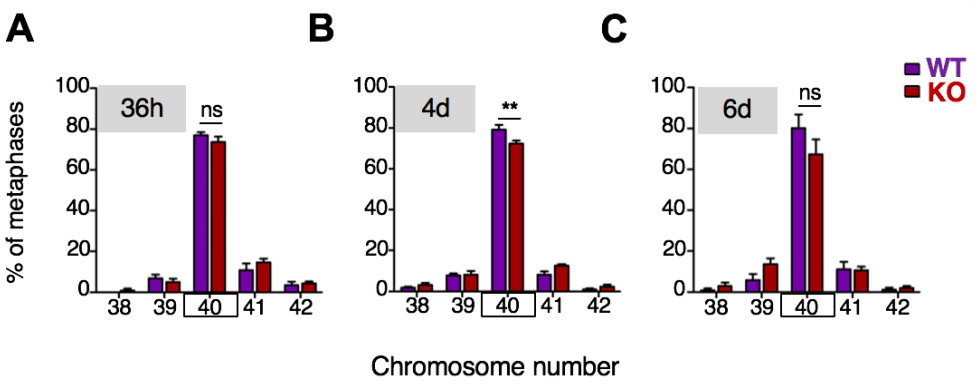


Figure R8. Reduced fidelity of chromosome segregation in SA2 depleted MEFs.
A-C. Quantification of chromosome number frequency in metaphase spreads of WT and KO MEFs (mean ± SEM). The different timepoints refer to the time KO cells have been cycling in the presence of 4-OHT. In A, 36h refer to 3 days of 4-OHT treatment upon serum starvation, followed by 36h of release from the arrest (as in Fig. R6A). In B and C,4 and 6 days indicate number of days in the presence of 4-OHT of asynchronously growing cells, respectively. For each condition, at least 100 metaphases from MEFs from 3 different embryos were inspected. Mann-Whitney test; *** P<0.001, ** P<0.01, * P<0.05, ns P≥0.05.

3. Specific roles of cohesin-SA2 in genome organization and gene regulation

In addition to its canonical function in sister chromatid cohesion, cohesin is essential for genome organization in interphase. It does so by stabilizing cis chromatin loops of different range, mostly together with CTCF but also independently of this architectural protein. This section explores the specific roles of cohesin-SA2 in genome organization and their impact on gene regulation. For that, we have examined genome wide distribution of cohesin and the transcriptomes of WT and SA2 KO MEFs, as well as of SA1 KO MEFs, previously generated in our group. We have also compared the chromatin association dynamics of the two cohesin variants by live cell imaging.

3.1 Different genome-wide distribution of cohesin-SA1 and cohesin-SA2

To explore genomic distribution of both cohesin variants, chromatin immunoprecipitation followed by deep sequencing (ChIP-seq) with custom-made, validated antibodies against cohesin subunits was used. In particular, ChIP-seq was carried our for SA2 in WT and SA1 KO MEFs and for SMC1 in WT and SA2 KO MEFs. Additional datasets were obtained from published and unpublished studies from our own group (Remeseiro *et al.* 2012, Ana Cuadrado and Miguel Ruiz Torres Doctoral Thesis UAM 2017) or other groups (Busslinger *et al.* 2017; Tedeschi *et al.* 2013)(Tedeschi 2013, Busslinger 2017). Reads were aligned to mm9 version of the mouse genome, peaks were called using MACS2 with FDR <0.05 and data were visualized by creating read density heatmaps.

First of all, we analyzed positions of SMC1 (which identify all cohesin complexes) and CTCF in WT MEFs and encountered two distinct populations of cohesin. A majority of cohesin positions (71%) overlapped with CTCF positions (75,071 CTCF-cohesin sites), while the other 29% did not (30,608 non-CTCF-cohesin sites, Fig. R9A). To explore whether both cohesin isoforms contributed to both types of positions, we layered these data with those obtained for SA1 and SA2 in WT MEFS. The two cohesin variants were robustly present at “CTCF-cohesin” sites, while cohesin-SA2 was predominant at non-CTCF-cohesin positions (Fig. R9A). Analyses of cohesin distribution in SA1KO and SA2KO MEFs revealed that non-CTCF-cohesin positions were retained in the absence of SA1 but lost in the absence of SA2, while CTCF-cohesin positions were firmly maintained in the absence of either variant (Fig. R9B).

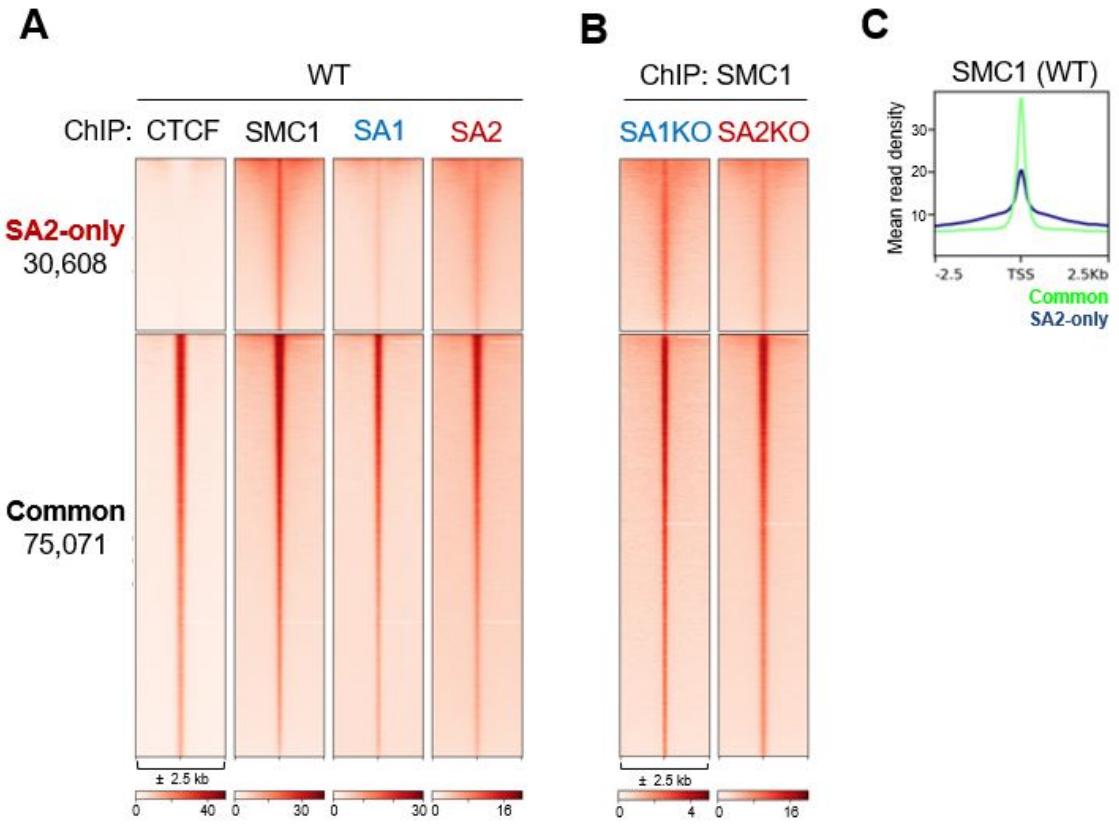


Figure R9. Two classes of genome-wide cohesin binding sites in MEFs.

A. Read density heatmaps showing ChIP-seq read distribution of CTCF, SMC1, SA1 and SA2 at common and SA2-only cohesin positions in WT MEFs. Reads are plotted in a 5 kb window around cohesin peak summits.
B. Read density heatmaps for SMC1 distribution at common and SA2-only positions (from A) in SA1KO and SA2KO MEFs.
C. Mead read density plot for SMC1 at common and SA2-only cohesin positions in WT MEFs.

We conclude that both cohesin-SA1 and cohesin-SA2 occupy sites bound by CTCF independently of one another while cohesin-SA2 is the preferred complex at positions not bound by CTCF. Moreover, cohesin-SA1 cannot occupy these non-CTCF sites even when cohesin-SA2 is absent. We therefore refer to the positions shared between the two variant complexes as “common” positions, and those specific for cohesin-SA2 as “SA2-only”-positions.

In a further effort to understand the nature and function of these two types of positions, we observed that common cohesin positions had higher and narrower peaks than SA2-only cohesin positions, as shown in the mean read density plot for SMC1 (Fig. R9C). This result suggests that SA2-only positions are more variable from cell to cell in the population and/or more transient than common positions. This idea urged us to explore possible differences in the association of cohesin variants with chromatin.

3.2 Cohesin-SA2 is less tightly bound to chromatin

As an initial proxy to explore the strength of the association to chromatin of the two variant cohesin complexes, we performed a biochemical assay in which we challenged chromatin-bound proteins with a high salt concentration. We first isolated the chromatin fraction using a biochemical fractionation protocol, then incubated this fraction with a buffer containing 0.5M NaCl for a short time and finally assessed by immunoblot the protein fraction still bound to chromatin after treatment.

First, we performed this assay in WT MEFs and observed that a larger amount of SA1 remained on chromatin after salt extraction compared to SA2 (lanes 1 and 2 in Fig. R10A). Next, we confirmed this result in SA1KO and SA2KO MEFs, where we focused on Rad21 levels for a direct comparison. In SA1KO MEFs, which contain only cohesin-SA2, a larger amount of Rad21 is extracted from chromatin upon salt treatment (lanes 3 and 4 in Fig. R10A). Only 19% of Rad21 present in untreated chromatin remained bound to chromatin (Fig. R10B). In SA2KO MEFs, where the only remaining cohesin is cohesin-SA1, the release of Rad21 is less pronounced (lanes 5 and 6 in Fig. R10A) and 62% of the protein present in unchallenged chromatin is still bound after salt incubation (Fig. R10B). In WT MEFs, an intermediate amount of Rad21 (42%) resists the salt extraction (lanes 1 and 2 in Fig. R10A; Fig. R10B), consistent with the presence of both cohesin variants.

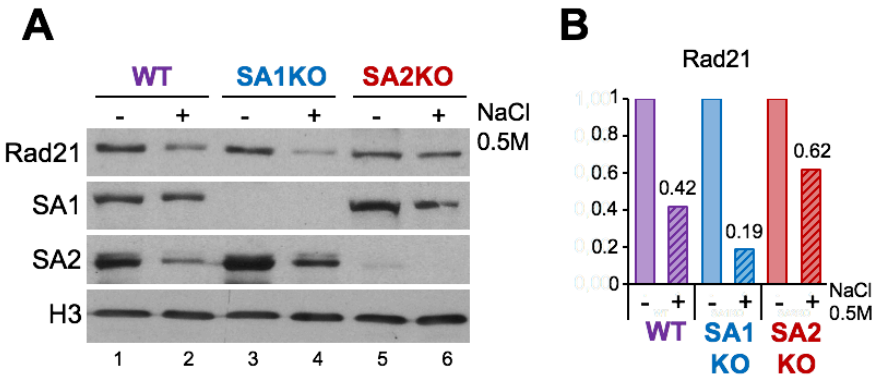


Figure R10. Cohesin-SA2 is more easily extracted from chromatin upon salt treatment.

A. Immunoblot analysis of chromatin fractions untreated or treated with 0.5M NaCl in MEFs of different genotype (WT, SA1KO and SA2KO).

B. Quantification of Rad21 from A. Numbers above the bars indicate % of remaining cohesin, normalized to H3 levels and relative to the untreated condition in each case.

Both results in WT and KO MEFs show that cohesin-SA2 is more sensitive to salt extraction than cohesin-SA1 and thus provide initial evidence to propose that cohesin-SA2 is less strongly bound to chromatin.

3.3 iFRAP: an assay to measure cohesin dynamics

A more refined method to address the dynamics of cohesin association to chromatin is a technique known as inverse Fluorescence Recovery After Photobleaching (iFRAP). This is a modification of the classical FRAP technique, which is used to evaluate mobility of fluorescently tagged proteins by photobleaching a small area within the cell and monitoring the recovery of fluorescence in that area with time-lapse imaging (Fig. R11A). iFRAP is typically applied for chromatin-bound proteins that have longer recovery times compared to soluble proteins in diffusion. In iFRAP, the bleached area is larger and monitoring times are longer than for soluble proteins (minutes vs. seconds). The rate of recovery is estimated by taking into account the redistribution of fluorescence in both the bleached area and the unbleached area (Fig. R11B). Since cohesin is mostly a chromatin-bound complex, iFRAP was a suitable technique to assess the dynamics of cohesin variants.

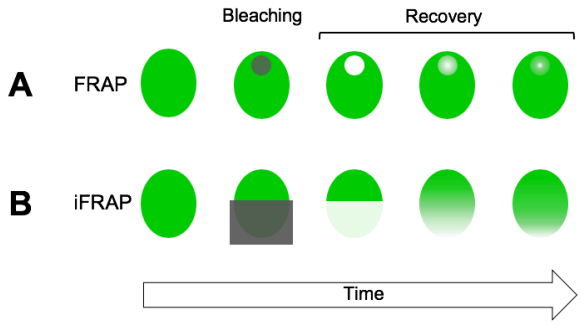


Figure R11. Schematic representation of FRAP and iFRAP experiments.
See text for details.

3.4 Generation of cell lines with GFP-tagged cohesin subunits for iFRAP

For iFRAP we needed to generate MEFs expressing cohesin subunits fused to a fluorescent tag. To ensure physiological expression of the tagged proteins, we performed genome editing of endogenous genes with CRISPR-Cas9. Unlike the rest of experiments in this thesis, we used MEFS that were immortalized by transduction with the large T antigen of the SV40 virus (iMEFs). Using the same parental WT iMEFs, we generated cell lines with GFP tags at the C-terminus of SA1, SA2 and Rad21 cohesin subunits.

Instead of the classical system (Fig. R12A), we chose a modified system that relies on a mutated version of Cas9 (Cas9n-D10A; Fig. R12B) that, once it reaches its target guided by a small guide RNA (sgRNA), introduces a nick in the DNA instead of a double strand break (DSB) (Trevino & Zhang 2014). For the system to work, two sgRNAs direct the

Cas9 to introduce two nicks in opposite strands, leaving a 5' overhang that is then repaired by HR (Fig. R12C). The need for two sgRNAs with limited offset reduces the off-target effects of this system. Homology arms that, in our case, target the C-terminus of each gene of interest, including an in-frame insertion of GFP right before the STOP codon are provided to be used as template for HR-repair. Paired sgRNA plasmids, along with a plasmid expressing the Cas9 nickase and the donor plasmid were electroporated into iMEFs. Cells that had substituted the WT allele with the GFP-fused version were selected by cell sorting based on the GFP signal and polyclonal populations were grown and used for our studies. We proceeded to validate these cell lines by immunoblot, immunofluorescence, chromatin fractionation and immunoprecipitation before performing iFRAP (Fig. R12D).

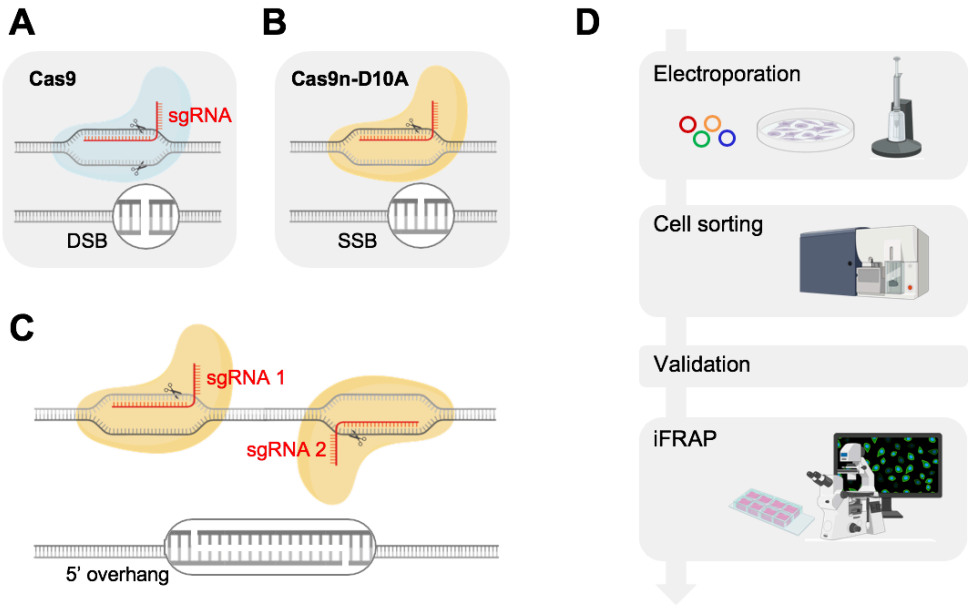


Figure R12. Generation of GFP-tagged cell lines by CRISPR targeting.

A. Wild type version of Cas9, that generates a DSB.
B. Mutated nickase version of Cas9, Cas9n-D10A, that generates a nick.
C. Double nickase system driven by two sgRNAs against opposite strands, generating a break with 5' overhang.
D. Schematic representation of the generation of GFP-tagged cell lines by electroporation until the iFRAP, the final assay.

3.5 Validation of GFP-tagged cell lines

To validate the successful targeting of the three cohesin subunits, we first looked at their GFP and cohesin levels by immunofluorescence in fixed cells. We could confirm the presence of cells positive for GFP. The GFP signal coincides with cohesin (SMC1) staining, both in localization to the nucleus and in its pattern (Fig. R13A). However, GFP-

negative cells were also seen. As opposed to systems relying on overexpression, here GFP expression was limited to the endogenous expression of the targeted genes. This, combined with limited efficiency of co-transfection, rendered negative cells in the sorted population. However, this was not a major issue for iFRAP because this is a single-cell study. Since no difference in cohesin (SMC1) levels were detected between GFP-negative and GFP-positive cells (Fig. R13A), we concluded that GFP tagging did not significantly interfere with expression of the subunits.

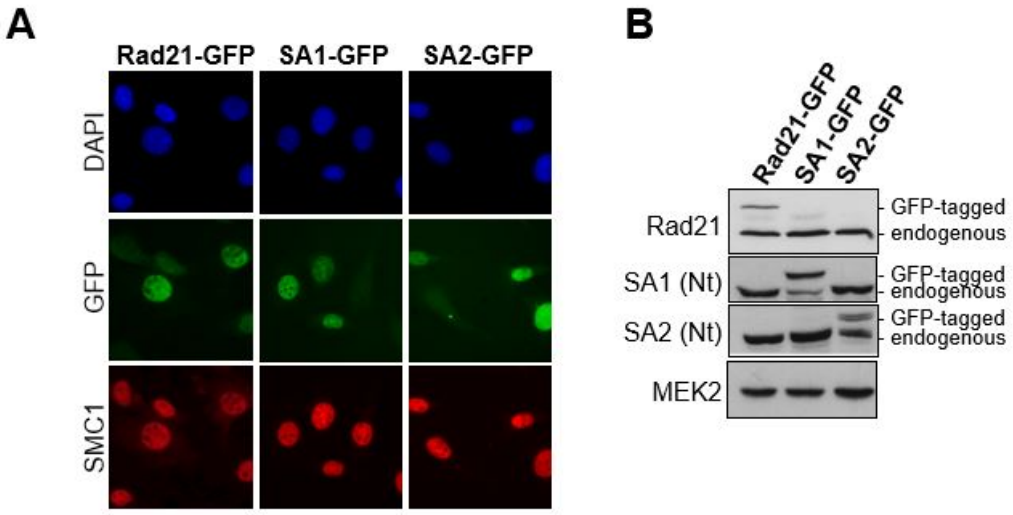


Figure R13. Successful generation of cells with GFP-tagged cohesin subunits.

A. Images from fluorescent microscope. GFP-tagged cells were stained with SMC1 antibody and DAPI to counterstain DNA, whereas GFP signal was visualized directly.
B. Immunoblot of whole-cell extracts from Rad21-, SA1- and SA2-GFP tagged polyclonal populations. MEK2 is used as a loading control.

We next performed immunoblotting of total cell extracts of the three different cell lines. For Rad21-GFP, we observe the endogenous Rad21 band, along with a band that migrates at a higher molecular weight, corresponding to Rad21-GFP. The presence of both bands in the population can be explained because targeting most likely occurs in heterozygosity, with only one allele being tagged with GFP and coexisting with the WT protein. The endogenous band is more prominent, likely due to the presence of GFP-negative cells in the population (lane 1 in Fig. R13B). For SA1 and SA2 we used newly generated antibodies against the N-terminal parts of the protein (Nt), which allowed correct detection of the C-terminal GFP-tagged proteins. For SA1-GFP, we could again detect both the endogenous and tagged band. The tagged band is more prominent, suggesting that the targeting was quite efficient (lane 2 in Fig. R13B). For SA2-GFP, we also detected bands for both endogenous and GFP-tagged versions. The gene encoding SA2 is located on the X chromosome, and since the MEFs come from a male embryo (XY), the lower and upper bands corresponding to untagged and tagged SA2 come from unedited and edited cells, respectively (lane 3 in Fig. R13B).

Chromatin fractionation of MEFs expressing SA1-GFP or SA2-GFP followed by immunoblot showed that tagged SA1 and SA2 proteins were present in the chromatin fraction and not in the soluble fractions, just like the endogenous proteins (Fig. R14A). Moreover, the two tagged proteins were pulled down with an antibody against SMC1 (lanes 3 and 6 in Fig. R14B), suggesting they can form part of a cohesin ring. For the Rad21-GFP cell line we had previously obtained similar results (Miguel Ruiz Torres, Doctoral Thesis, UAM 2017).

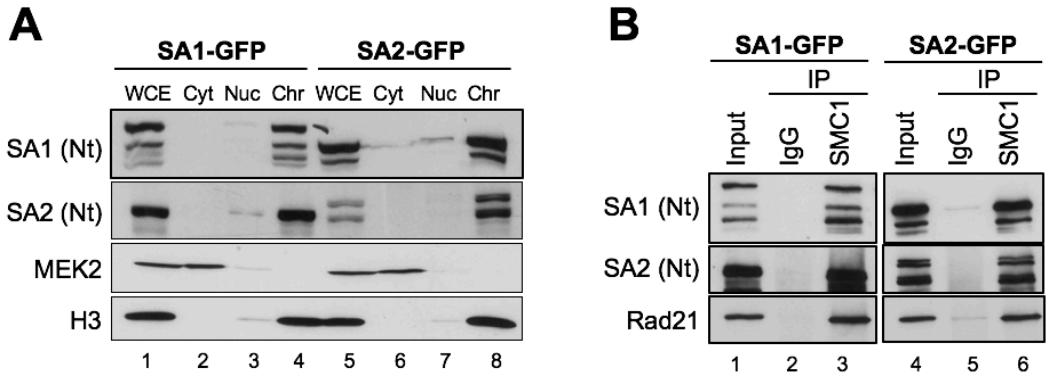


Figure R14. Validation of SA1- and SA2-GFP populations by chromatin fractionation and immunoprecipitation.

A. Immunoblot of cellular fractions from SA1- and SA2-GFP tagged cells. Whole cell extract (WCE), cytosol (Cyt), nucleoplasm (Nuc) and chromatin-bound (Chr) fractions were loaded. MEK2 and H3 were used as controls for correct fractionation.
B. Immunoprecipitations of SMC1 in SA1- and SA2-GFP tagged cells. IP with IgG antibody was used as a negative control. 3% of input and 30% of IP were loaded.

Taking together these experiments, we have correctly generated cells with GFP-tagged SA1, SA2 and Rad21 subunits that are able to target chromatin and do so in the context of a cohesin complex.

3.6 Different dynamic behavior of SA1 and SA2 by iFRAP

During S phase, a “cohesive” cohesin population is generated to hold the sister chromatids together that is not sensitive to WAPL-mediated release and remains stably bound to chromatin until mitosis. To avoid this population and restrict our analyses to the dynamic cohesin population, we performed iFRAP experiments in cells arrested in G0. For this, cells were seeded at high confluence in a microslide plate and kept in low serum medium (0.1% FBS) for 1 day before taking them to the confocal microscope. A high-potency laser was used to photobleach half of the nucleus of selected cells and fluorescence recovery was monitored with interval captures over the following 20 minutes.

For each timepoint, the difference in mean fluorescence intensity between the bleached and unbleached areas relative to the initial difference at the time of bleaching was calculated. Results were normalized both to total fluorescence intensity of the nucleus and to the background signal. Average data were fitted to a single exponential curve.

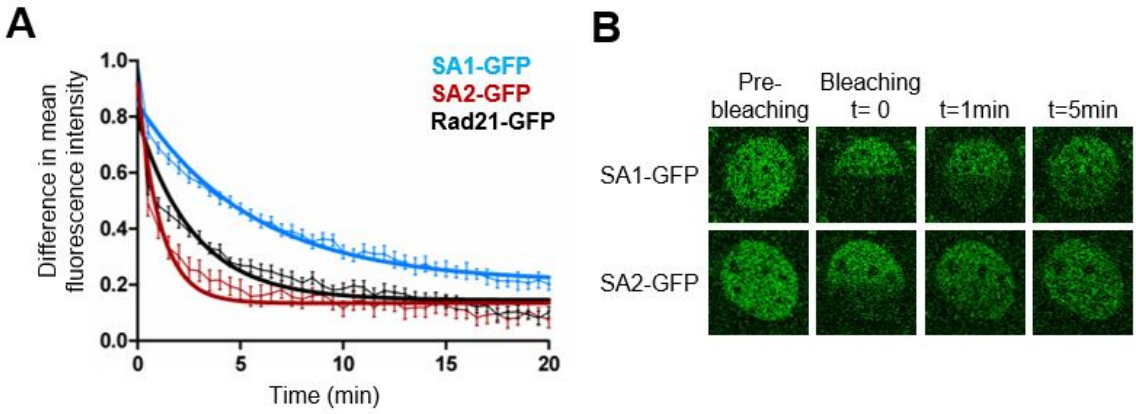


Figure R15. iFRAP reveals distinct dynamics of cohesin-SA1 and cohesin-SA2.

A. Plot of fluorescence recovery in SA1-GFP, SA2-GFP and Rad21-GFP cell lines, representing the difference in mean fluorescence intensity between the bleached and the unbleached areas relative to the time of bleaching. Mean ± SEM are shown (thin line) and a single exponential curve fit has been applied (thick line). n=20 for Rad21-GFP, n=24 for SA1-GFP and n=16 cells for SA2-GFP, from 4 independent experiments.

B. Representative image sequences of iFRAP in SA1- and SA2-GFP cells, showing pre-bleaching, post-bleaching and recovery frames.

We observed that the recovery of fluorescence in the bleached area was faster in SA2-GFP cells than in SA1-GFP cells, while Rad21-GFP behavior was in between both (Fig. R15A). Consistently, the residence times calculated in Graphpad were 1.08 minutes for SA2 and 5.56 minutes for SA1. The difference was clear at early timepoints (up to 10 minutes) and all cell lines reached a similar final recovery at 20 minutes. Representative images are shown (Fig. R15B).

This result indeed suggests that cohesin-SA2 is more dynamic in its association to chromatin than cohesin-SA1, with an average residence time on chromatin that is 5-fold shorter.

3.7 Preferential interaction of cohesin-SA2 with cohesin release factor WAPL

To find a molecular explanation as to why cohesin variants presented this different dynamic behavior, we looked into differential interactions with one of the key proteins that modulate cohesin binding to chromatin: WAPL. We found that the amount of WAPL present in immunoprecipitates of cohesin obtained with an antibody against SMC1 was lower in SA2-null MEFs than in WT MEF extracts (Fig. R16). This result suggests that there is a preferential interaction of cohesin-SA2 complexes with WAPL.

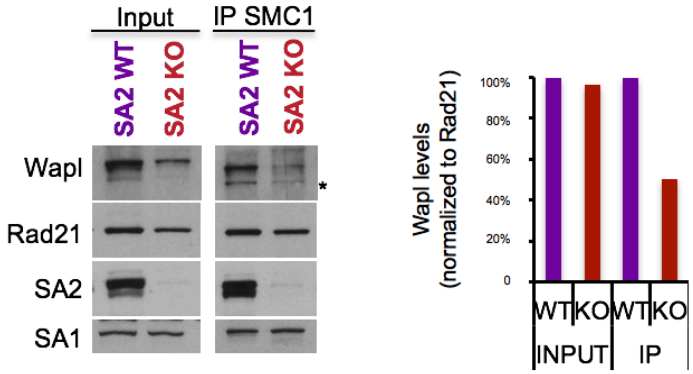


Figure R16. Preferential interaction of cohesin-SA2 with WAPL.
Left, immunoblot analysis of SMC1 immunoprecipitates in SA2 WT and KO MEFs. Asterisk indicates a non-specific band in the WAPL immunoblot. Right, quantification of WAPL levels in input and immunoprecipitated relative to Rad21.

With this new information, we wondered how depletion of WAPL would specifically affect the genome-wide positioning of cohesin, especially at SA2-only sites. A previous study had shown that WAPLKO MEFs had increased amount of cohesin on chromatin, that this cohesin was stably bound and that its genome wide distribution was similar to that of WT MEFs (Tedeschi *et al.* 2013). However, our reanalysis of these ChIP-seq data revealed that cohesin (SMC3) is lost specifically in SA2-only positions relative to common positions in the absence of WAPL (Fig. R17A). As these SA2-only positions are not detected when cohesin is stabilized on chromatin, they most likely depend on the dynamic behavior of cohesin-SA2.

Cohesin-SA1, on the other hand, is less dynamic and, in WT cells, accumulates only in CTCF-bound sites. Reanalysis of ChIP-seq data from CTCFKO MEFs (Busslinger *et al.* 2017) showed that the presence of cohesin-SA1 at common positions was drastically reduced in the absence of CTCF and instead the complex occupies SA2-only positions (Fig. R17B). A recent preprint reports that CTCF depletion in HeLa cells decreases the

stability of cohesin-SA1 binding to chromatin. Although we have failed to see an effect of CTCF knock down on the salt extractability of SA1 (data not shown), an attractive hypothesis is that stabilization of cohesin-SA1 at CTCF sites may render this variant less dynamic and unable to occupy the more dynamic SA2-only sites.

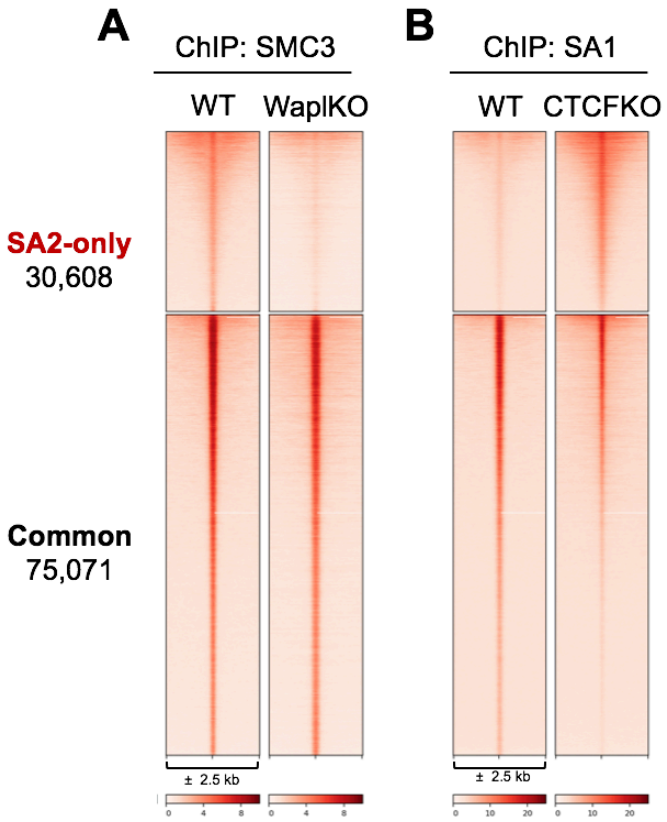


Figure R17. Cohesin redistribution upon depletion of WAPL and CTCF.

A. Read density heatmaps showing ChIP-seq read distribution of SMC3 at common and SA2-only cohesin positions in WT and WAPLKO MEFs. Reads are plotted in a 5 kb window around cohesin peak summits. Data obtained from Tedeschi *et al.* 2013 and reanalyzed.

B. Read density heatmaps of SA1 distribution at common and SA2-only cohesin positions in WT and CTCFKO MEFs. Data obtained from Busslinger *et al.* 2017 and reanalyzed.

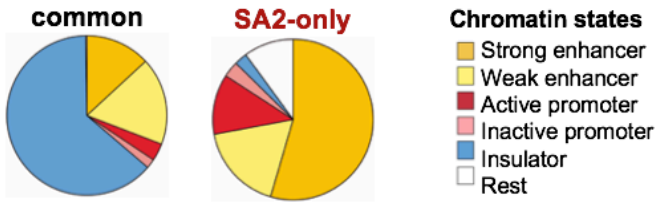
Taken together, we propose that SA2-only sites are dependent on the dynamic behavior of cohesin-SA2 and that cohesin-SA1 may occupy these sites when not being retained by CTCF.

3.8 SA2-only positions are enriched in active chromatin states and transcription factor binding motifs

For further insight into the relevance of cohesin distribution in common and SA2-only sites, we asked about additional genomic features of these positions using chromatin states defined in MEFs. Common cohesin positions are mostly enriched in insulators

defined by the presence of CTCF and no other histone modification (pie chart on the left in Fig. R18A). In contrast, SA2-only positions are found mostly at enhancers and promoters, especially active ones (pie chart on the right in Fig. R18A). The exclusive presence of cohesin-SA2 at a subset of cis-regulatory elements suggests that this complex may have a non-redundant role in the regulation of gene expression.

A



B

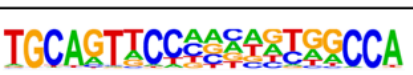
Positions	Name	Motif	p-value	Enrichment over background
Common	CTCF		1E-38455	61.12 vs 1.52 %
	CTCFL		1E-34782	64.46 vs 2.58 %
	CTCF satellite element		1E-2932	5.63 vs 2.58%
	THRb		1E-1592	15.84 vs 5.10 %
	COUP		1E-1327	30.46 vs 15.94 %
SA2-only	Fra1		1E-3385	24.08 vs 3.99%
	Atf3		1E-3277	26.08 vs 4.98 %
	BATF		1E-3181	25.72 vs 3.99 %
	JunB		1E-3178	23.78 % vs 4.18 %
	Fra2		1E-3147	22.27 vs 3.63 %

Figure R18. Chromatin states and motif analysis in common and SA2-only cohesin positions.

A. Pie charts showing distribution of common and SA2-only positions in chromatin states defined by epigenetic marks in MEFs (ENCODE), defined as follows: H3K27ac and H3K4me1 for strong enhancers, H3K4me1 for weak enhancers, H3K27ac and H3K4me3 for strong promoters, H3K4me3 for weak promoters, CTCF for insulators, or none of the above for “rest”.

B. Top 5 motifs enriched in common and SA2-only cohesin positions by Homer Known Motif Enrichment Analysis. Enrichment of the motif in the target sequences over the presence of the same motif in an equal number of randomly chosen sequences (background) is shown.

Motif enrichment analysis using HOMER in common and SA2-only cohesin positions also revealed differences in the binding sequences. In common positions the top enriched motif was CTCF, present in 61% of positions, closely followed by CTCFL/BORIS, its germline paralog. Motifs for additional transcription factors were also identified (Fig. R18B). In the SA2-only cohesin positions there is a notable enrichment of binding motifs for transcription factors from the AP-1 family (Fra1, Fra2, Atf3, BATF, JunB) and also transcription factors belonging to the TEAD family.

The exclusive enrichment of certain transcription factor binding motifs in SA2-only positions suggest an important contribution of cohesin-SA2 present at these positions to transcriptional control independently of cohesin-SA1.

3.9 Cells depleted of SA1 and SA2 have a different effect on gene expression

To evaluate the specific contributions of cohesin-SA1 and cohesin-SA2 to gene regulation, we performed RNA-sequencing in 3 pairs of WT and *Stag1* KO MEFs as well as 3 clones of *Stag2* cKO MEFs treated or not with 4-OHT for 4 days. A total of 917 differentially expressed genes (DEGs, FDR<0.05; fpkm>2 in at least one condition) were identified in SA1KO MEFs (370 upregulated, 547 downregulated, average log_2FC 1.2 and -2.1, respectively, range +6.5 to -6.2). A reduced number of DEGs, 246, was found after elimination of SA2 (65 genes upregulated, 181 downregulated, average log_2FC 0.85 and -0.93, respectively, range +2.4 to -3.3). The more extensive transcriptional deregulation of SA1KO MEFs can be in part due to the fact that we are comparing MEFs derived from different embryos, whereas to test the effect of SA2 elimination, MEFs are derived from the same embryo untreated (WT) or 4-OHT-treated (KO).

Gene Set Enrichment Analysis (GSEA) uncovered radically different profiles of gene deregulation in the absence of either cohesin variant (Fig. R19). In fact, many significantly deregulated pathways (FDR<0.05), were altered in opposite directions (Fig. R19 and R20A). For instance, several pathways related to cell cycle and DNA repair were downregulated in SA2KO and upregulated in SA1KO MEFs (Fig. R20B). Pathways displaying the opposite trend, i.e. upregulated in SA2 KO and downregulated in SA1 KO MEFs were membrane-related processes including “Lysosome” or receptor signaling (Fig. R20C).

GSEA in SA2KO			
Downregulated (46)	**FDR**	**Upregulated (16)**	**FDR**
RIBOSOME	0	LYSOSOME	0
SPLICEOSOME	0	ENDOCYTOSIS	9.68E-04
DNA REPLICATION	0	REGULATION OF AUTOPHAGY	0.0029
RNA TRANSPORT	0	UBIQUITIN MEDIATED PROTEOLYSIS	0.0024
FANCONI ANEMIA PATHWAY	0	SNARE INTERACTIONS IN VESICULAR TRANSPORT	0.0048
BASE EXCISION REPAIR	0	PEROXISOME	0.0047
MISMATCH REPAIR	0	OTHER GLYCAN DEGRADATION	0.0079
ARRHYTHMOGENIC RIGHT VENTRICULAR CARDIOMYOPATHY (ARVC)	0	INOSITOL PHOSPHATE METABOLISM	0.0146
ECM-RECEPTOR INTERACTION	0	COLLECTING DUCT ACID SECRETION	0.0136
CELL CYCLE	0	FOXO SIGNALING PATHWAY	0.0166
AMINOACYL-TRNA BIOSYNTHESIS	0	B CELL RECEPTOR SIGNALING PATHWAY	0.0153
RIBOSOME BIOGENESIS IN EUKARYOTES	0	SYNAPTIC VESICLE CYCLE	0.0220
HOMOLOGOUS RECOMBINATION	7.97E-05	GLYCOSAMINOGLYCAN DEGRADATION	0.0204
RNA DEGRADATION	4.57E-04	PHOSPHATIDYLINOSITOL SIGNALING SYSTEM	0.0214
PROTEOGLYCANS IN CANCER	4.26E-04	OLFACTORY TRANSDUCTION	0.0276

GSEA in SA1KO			
Downregulated (35)	**FDR**	**Upregulated (40)**	**FDR**
LYSOSOME	0	SPLICEOSOME	0
PHAGOSOME	6.72E-04	RNA TRANSPORT	0
OTHER GLYCAN DEGRADATION	4.48E-04	DNA REPLICATION	0
STAPHYLOCOCCUS AUREUS INFECTION	3.36E-04	RIBOSOME BIOGENESIS IN EUKARYOTES	0
GLYCOSAMINOGLYCAN DEGRADATION	2.69E-04	RIBOSOME	0
NATURAL KILLER CELL MEDIATED CYTOTOXICITY	2.24E-04	CELL CYCLE	0
RHEUMATOID ARTHRITIS	1.92E-04	FANCONI ANEMIA PATHWAY	0
OSTEOCLAST DIFFERENTIATION	6.6E-04	MRNA SURVEILLANCE PATHWAY	0
B CELL RECEPTOR SIGNALING PATHWAY	5.86E-04	MISMATCH REPAIR	0
TOLL-LIKE RECEPTOR SIGNALING PATHWAY	5.28E-04	PYRIMIDINE METABOLISM	0
TUBERCULOSIS	4.80E-04	PROTEASOME	0
INFLUENZA A	6.63E-04	RNA POLYMERASE	0
VALINE, LEUCINE AND ISOLEUCINE DEGRADATION	0.0014	PURINE METABOLISM	0
MEASLES	0.0013	HOMOLOGOUS RECOMBINATION	0
CYTOKINE-CYTOKINE RECEPTOR INTERACTION	0.0019	BASE EXCISION REPAIR	0

Figure R19. GSEA analysis of transcriptomes from SA1KO and SA2KO MEFs.
Top 15 upregulated and downregulated KEGG pathways are shown for each condition out of all significantly deregulated pathways at FDR<0.05. The total number of significant pathways is indicated between brackets.

Enrichment profiles for some pathways are displayed as examples and show the striking opposite effect of the absence of one or the other variant cohesin complex (Fig. R20B and R20C). The opposite transcriptional signatures observed between SA1KO and SA2KO MEFs points at important implications for both cohesin variants in gene regulation, but suggest that the mechanisms or targets of this control are different for cohesin-SA1 and cohesin-SA2.

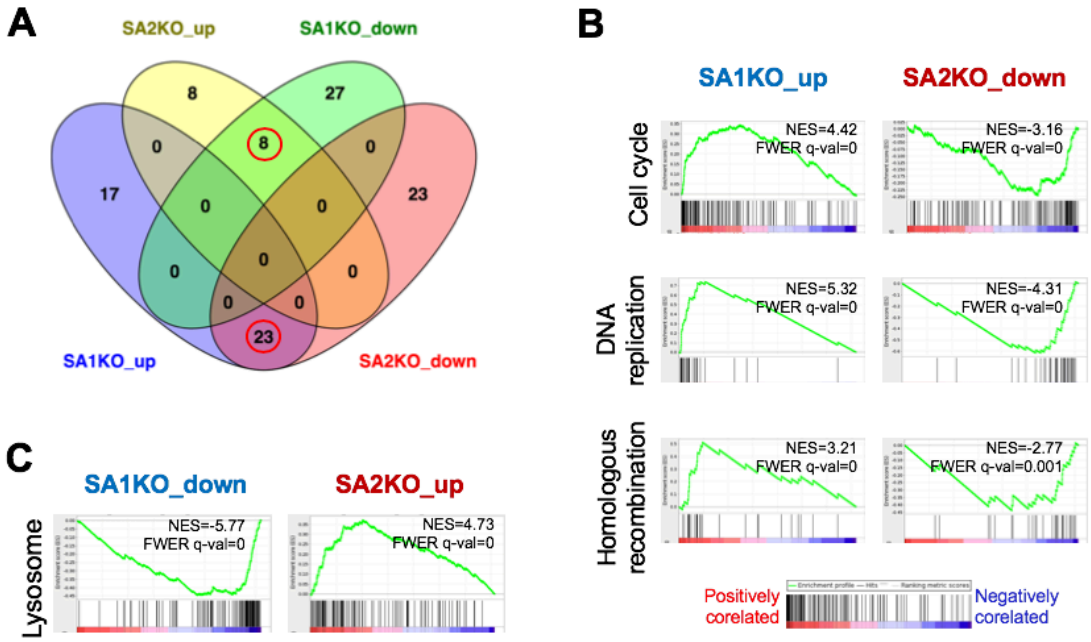


Figure R20. Cells depleted of SA1 and SA2 have an opposite transcriptional signature.

A. Venn diagram showing the overlap between KEGG pathways significantly up- or downregulated at FDR<0.05 in SA1KO and SA2KO MEFs. Data from GSEA.

B-C. Enrichment plots of representative GSEA pathways that are upregulated in SA1KO and downregulated in SA2KO MEFs (B) or viceversa (C). Normalized enrichment score (NES) and FWER q-value are shown in each plot.

We hypothesize that this different control of gene expression by cohesin-SA1 and cohesin-SA2 could be again related to their distinct dynamics of binding to chromatin. To further test this idea, we compared the transcriptomic changes described above with those previously obtained in MEFs lacking cohesin releasing factors Pds5A and Pds5B (Figure 21A). FRAP analyses in these cells had revealed a clear stabilization of cohesin dynamics in the absence of both Pds5 proteins (Pds5KO, Miguel Ruiz Torres Doctoral Thesis UAM 2017 and Morales *et al.* 2019). We found a significant overlap between the pathways downregulated in Pds5-deficient MEFs and those downregulated in SA2KO MEFs and upregulated in SA1KO MEFs (49 and 52%, respectively). Similarly, pathways upregulated in Pds5KOMEFs overlapped with pathways upregulated SA2-null MEFs and downregulated in SA1KO MEFs (27% and 42%) (Fig. R2B). Thus, stabilization of cohesin binding to chromatin by reducing Pds5 protein levels or by removing the more dynamic cohesin, cohesin-SA2, results in comparable gene expression changes.

A

GSEA in Pds5KO			
Downregulated (39)	**FDR**	**Upregulated (33)**	**FDR**
CELL CYCLE	0	LYSOSOME	0
RIBOSOME	0	RHEUMATOID ARTHRITIS	0
SPLICEOSOME	0	PHOTOTRANSDUCTION	0
RNA TRANSPORT	0	CYTOKINE-CYTOKINE RECEPTOR INTERACTION	0
FANCONI ANEMIA PATHWAY	0	SYNAPTIC VESICLE CYCLE	0
DNA REPLICATION	0	RETINOL METABOLISM	1.81E-04
RIBOSOME BIOGENESIS IN EUKARYOTES	0	OXIDATIVE PHOSPHORYLATION	3.19E-04
RNA DEGRADATION	0	TYPE I DIABETES MELLITUS	0.0019
ARRHYTHMOGENIC RIGHT VENTRICULAR CARDIOMYOPATHY (ARVC)	0	NON-ALCOHOLIC FATTY LIVER DISEASE (NAFLD)	0.0020
DILATED CARDIOMYOPATHY	0	PHAGOSOME	0.0031
BASE EXCISION REPAIR	0	COLLECTING DUCT ACID SECRETION	0.0031
HYPERTROPHIC CARDIOMYOPATHY (HCM)	0	NEUROACTIVE LIGAND-RECEPTOR INTERACTION	0.0035
MRNA SURVEILLANCE PATHWAY	6.54E-05	GRAFT-VERSUS-HOST DISEASE	0.0032
PURINE METABOLISM	6.07E-05	PARKINSON'S DISEASE	0.0046
OOCYTE MEIOSIS	5.67E-05	STAPHYLOCOCCUS AUREUS INFECTION	0.0065

B

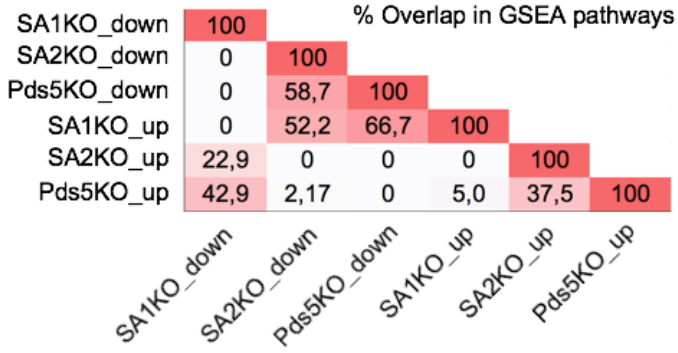


Figure R21. The transcriptome of Pds5KO MEFs resembles that of SA2KO MEFs.

A. Top 15 upregulated and downregulated KEGG pathways in Pds5KO MEFs out of all significantly deregulated pathways at FDR<0.05. The total number of significant pathways is indicated between brackets. Data from Miguel Ruiz Torres, Doctoral Thesis, UAM 2017.

B. Plot indicating the percentage of overlap between upregulated and downregulated pathways in all three conditions (SA1KO, SA2KO and Pds5KO MEFs).

4. The relevance of cohesin-SA2 for embryonic development

In the previous sections we have shown that cohesin-SA2 has unique roles in cell proliferation, genome organization and gene regulation *in vitro*. The next section explores the importance of these roles in the context of a developing embryo.

4.1 SA2-null embryos die by E10.5

We first asked whether cohesin-SA2 is essential for embryonic development. To achieve ubiquitous deletion of *Stag2* from the zypote stage we used a *CAG-Cre* transgene expressing a Cre recombinase through a strong synthetic promoter (Belteki *et al.* 2005). SA2 conditional KO females ($Stag2^{lox/lox}$) were mated with males carrying a wild type *Stag2* allele and one copy of *CAG-Cre* ($Stag2^{+/Y}$; $CAG\text{-}Cre^{+/T}$). Since *Stag2* is a X-linked gene, male embryos in the offsping would be either wild type (WT, $Stag2^{lox/Y}$) or knock-out (KO, $Stag2^{\Delta/Y}$), while females would be WT ($Stag2^{lox/+}$) or heterozygous (HT, $Stag2^{\Delta/+}$), depending on inheritance of the *Cre* allele (Fig. R22A). Each of these four options has a probability of 25% according to normal mendelian inheritance. The viability of SA2-null embryos was assessed by extracting embryos from pregnant females at different embryonic stages, that were genotyped for the *Stag2* allele and the presence of the Y chromosome by PCR (Fig. R22B).

By E12.5 no SA2KO embryos were alive and few reabsorbed embryos were found. Embryos of the other genotypes were present at expected proportions (Fig. R22C). Three days earlier, by E9.5, alive SA2KO embryos were obtained at the expected frequency. Complete depletion of SA2 in these embryos was confirmed by immunofluorescence with an antibody against SA2 in sections of E9.5 WT and KO male embryos (Fig. 22D). Next, we extracted litters at E10.5 and observed that the vast majority of SA2KO embryos were dead, as evidenced by the lack of heartbeat (Fig. 22C). We conclude that in the absence of cohesin-SA2, embryonic lethality occurs between E9.5 and E10.5.

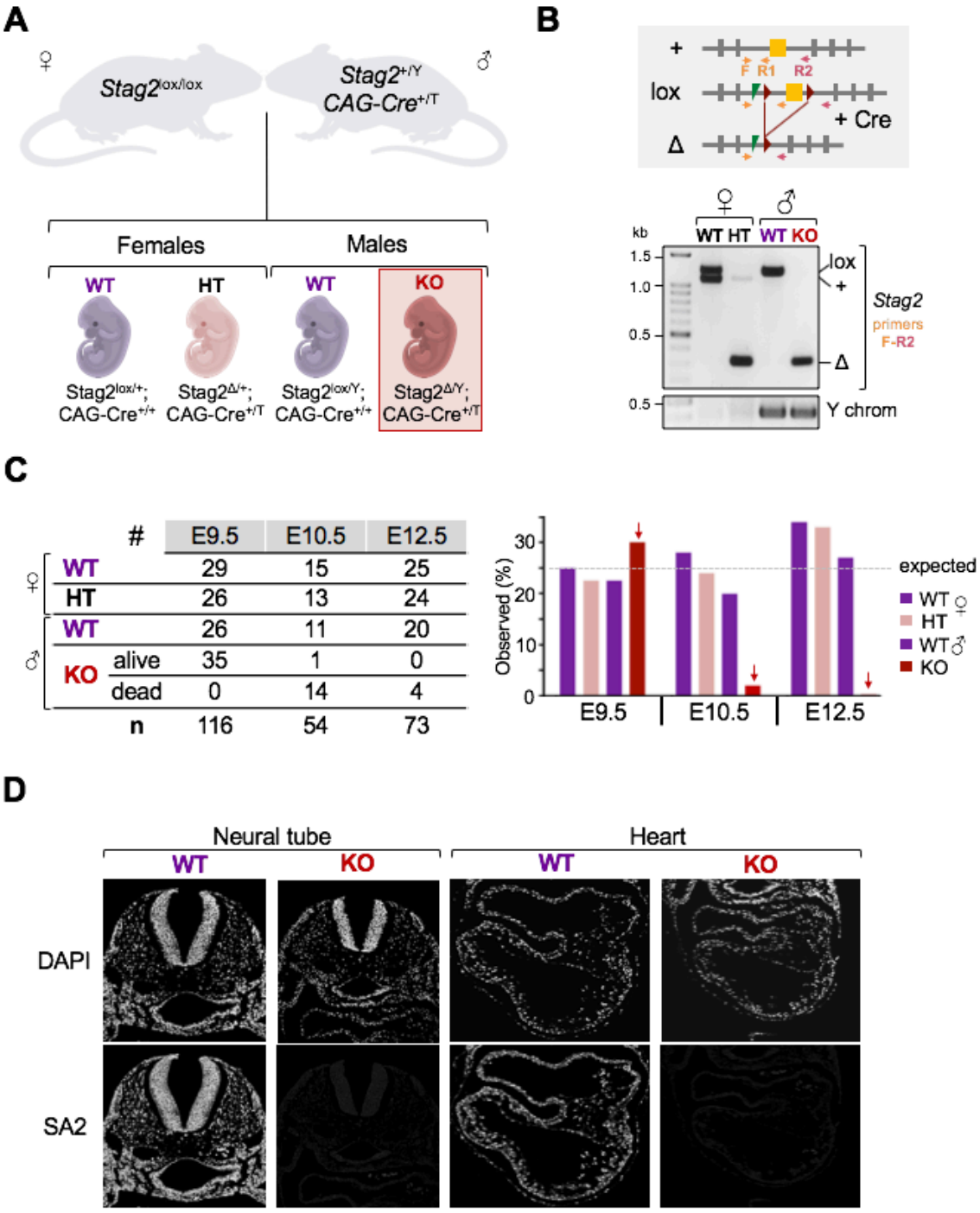


Figure R22. SA2-null embryos are embryonic lethal.

A. Mating strategy to obtain SA2KO embryos.

B. Strategy (top) and representative results (bottom) of embryo genotyping by PCR. Genomic DNA was extracted from embryo tissue or yolk sac. Amplification of the *Sry* gene was used to detect the Y chromosome.

C. Table (left) and bar graph (right) showing the proportions obtained for each of the genotypes in the offspring from the mating indicated in A at indicated stages. 14, 7 and 13 litters were obtained at E9.5, E10.5 and E12.5, respectively.

D. Immunofluorescence staining for SA2 of transverse sections of neural tube and heart of WT and SA2KO embryos at E9.5. Nuclei are counterstained with DAPI. Scale bar, 200 μm.

4.2 SA2 heterozygous embryos survive at submendelian proportions

With the mating strategy described above to obtain SA2KO male embryos, heterozygous females are also found in the offspring (HT, Fig. R22A). At mid-gestation they are present at near-normal ratios (HT, Fig. R22C). However, among the mice that reach birth from these crosses and are genotyped at the time of weaning, only half of the expected proportion of heterozygous female mice are obtained (HT, Fig. R23A). This suggests that a fraction of these females die between stage E12.5 and weaning. Those that survive appear as healthy as their WT female littermates (Fig. R23B). We hypothesize that they might have benefited from X-chromosome inactivation skewed towards the SA2 null allele. We continue exploring the phenotype of SA2 loss exclusively in male embryos.

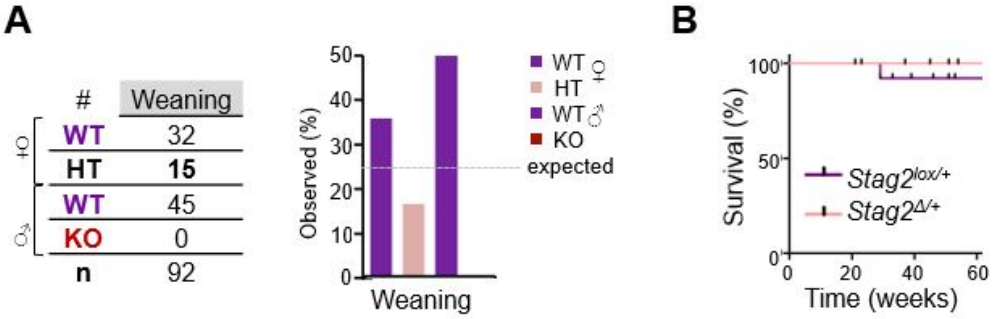


Figure R23. Half of SA2 heterozygous females survive until adulthood.

A. Table (left) and bar graph (right) showing genotypes obtained after weaning for the offspring of the mating in Fig. 22A. Data from 11 litters.
B. Kaplan-Meier survival plot for WT (*Stag2*$^{lox/+}$) and HT (*Stag2*$^{\Delta/+}$) females (n=15 mice of each genotype).

4.3 Lethality in SA2-null embryos is not caused by placental defects

A recent study has revealed the previously under-appreciated contribution of placental defects to early embryonic lethality (E9.5-14.5) of numerous mouse models (Perez-Garcia *et al.* 2018). In order to rule out that this was the case for SA2-null embryos, male mice carrying one copy of the *Sox2-Cre* transgene (*Stag2*$^{+/Y;}$ *Sox2-Cre*$^{+/T}$) were mated with SA2 conditional females (*Stag2*$^{lox/lox}$). The *Sox2* promoter drives *Cre* expression only in the epiblast thus ensuring that gene deletion does not take place in non-epiblast derived extraembryonic tissues, which remain WT (Hayashi *et al.* 2002). We observed that all SA2 KO embryos died by E12.5 (Fig. R24A), just as with the *CAG-Cre* allele. However, reabsorbed embryos found at this stage were a bit larger and had more integrity than the ones found at E12.5 after ubiquitous Cre expression, suggesting that embryo reabsorption and thus lethality happen slightly later (Fig. R24B). We can conclude that placental defects are not the main cause of lethality in SA2-null embryos

although they could contribute to accelerate it. Studies presented below were performed with SA2KO embryos carrying *CAG-Cre*.

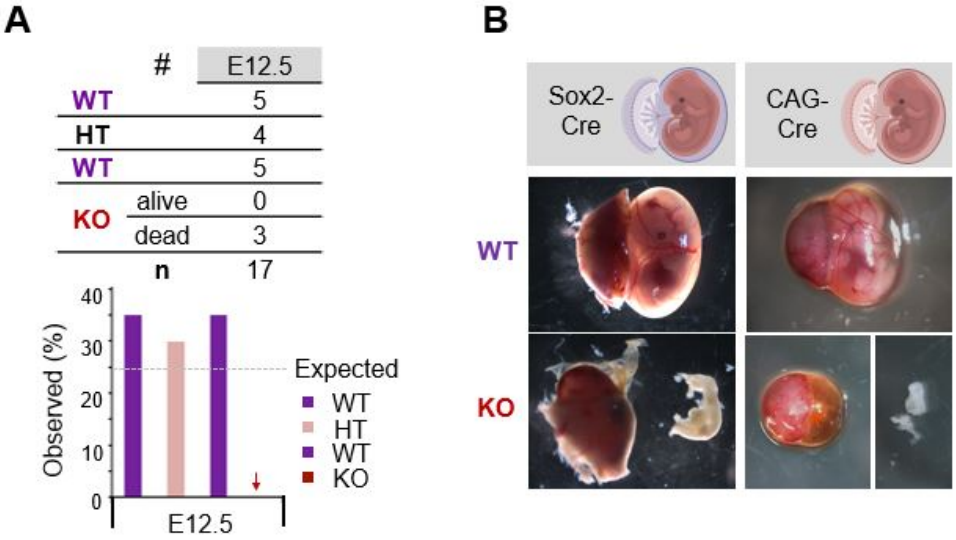


Figure R24. Lethality of SA2-null embryos is not due to extraembryonic defects.

A. Table (top) and bar graph (bottom) showing the outcome of 3 litters obtained by mating SA2 conditional females with males carrying a *Sox2-Cre* transgene Genotypes for *Stag2* are: female WT (lox/+), female HT (Δ/+), male WT (lox/Y), male KO (Δ/Y).
B. Representative comparison between WT and SA2KO male embryos at stage E12.5 using deletion mediated by either Sox2-Cre (Fig. R24A) or CAG-Cre (Fig. R22C). In the cartoons, red color indicates regions were Cre is activated and SA2 is depleted, as opposed to purple regions that remain WT.

4.4 SA2-null embryos display a developmental delay

To further explore the phenotype caused by SA2 depletion we inspected whole mount embryos. We focused our analysis on the male embryos (WT and SA2KO). At stage E9.5, prior to lethality, SA2KO embryos were already smaller in size than their WT littermates (Fig. R25A). We observed variability in the penetrance of the phenotype, as most SA2KO embryos did not display overt morphological defects (mild) but others presented alterations that went beyond the reduction in size (severe). These features were further enhanced at E10.5 (Fig. R25A, right).

To determine if the difference in size stemmed from a developmental delay, we counted pairs of somites in whole-mount embryos. Somites are segments of paraxial mesoderm that form bilaterally along the dorsal axis of developing vertebrate embryos (Fig. R25B). They give rise to vertebrae, ribs, skin and skeletal muscles, among other structures (Gilbert 2000). For mouse embryonic development, the number of somites expected at each stage is well established (Fig. R25C, right). Because individual embryos within the

same litter can develop at slightly different rates, the number of somites is a good indicator of actual developmental stage.

A significantly reduced number of somites was observed in mutant embryos starting at E9.5, indicating a clear developmental delay, while at E8.5 the somite counts were still similar. By E10.5 this difference was more pronounced and corresponded almost to a 1-day lag (Fig. R25C). Notably, there was no difference between the somite counts of SA2KO embryos of mild and severe phenotypes. Thus, loss of SA2 causes a generalized developmental delay, noticeable by E9.5 with variable penetrance.

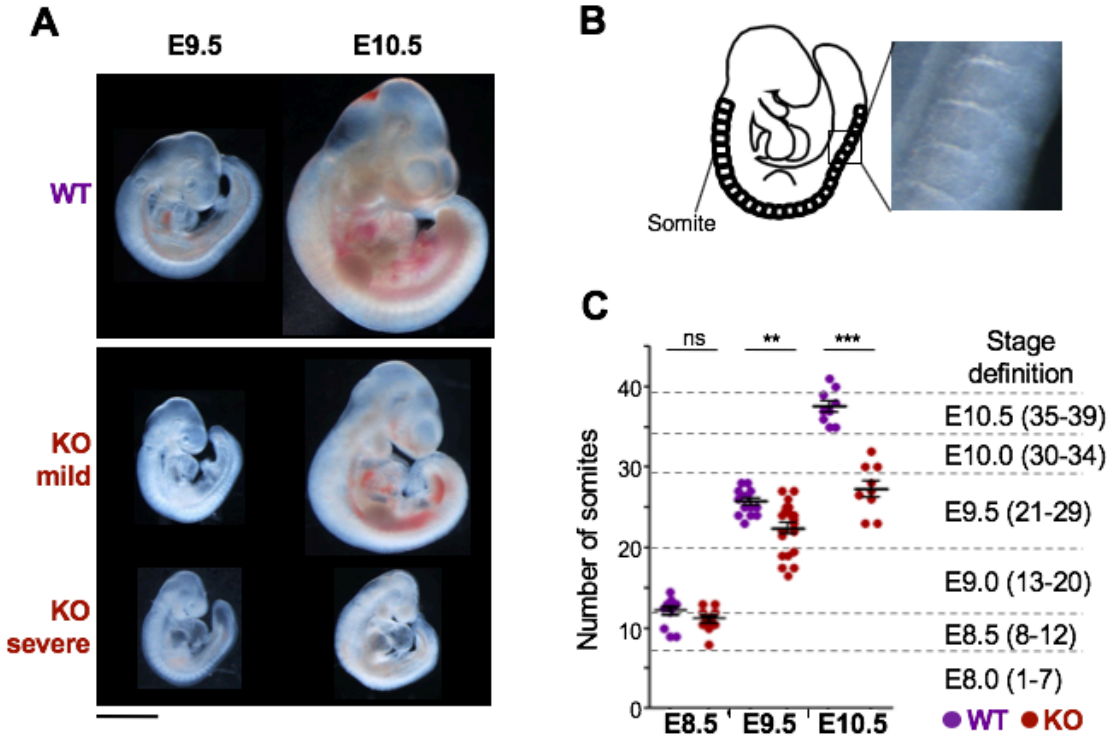


Figure R25. SA2-null embryos display a developmental delay starting at E9.5.

A. Representative images of WT (*Stag2*$^{lox/Y}$) and SA2KO (*Stag2*$^{\Delta/Y}$) male embryos at E9.5 and E10.5. SA2KO embryos can be classified into a mild and a severe phenotype. Scale bar, 1mm.

B. Schematic representation of somites in an E9.5 embryo (left) and a real image at high magnification where somites can be appreciated.

C. Somite number of WT and SA2KO embryos at indicated stages in embryos from 6 litters at E8.5 (n=13 WT and n=10 SA2KO embryos), 9 litters at E9.5 (n= 19 WT and n=25 SA2KO) and 4 litters at E10.5 (n=10 WT and n=10 SA2KO). Two-tailed Student's t-test, *** P<0.001, ** P<0.01, ns P≥0.05. At the right of the graph are the number of somites that correspond to each stage of murine development.

4.5 Global developmental defects in SA2-null embryos

We examined mutant embryos by histology to unveil defects underlying embryonic lethality. We proceeded with paired male WT and SA2KO littermates but included an additional control consisting of male WT embryos from a different litter with the same number of somites as the SA2KO embryos. We have termed the littermate or age-matched control embryos as WT1 and the somite- or stage-matched control embryos as WT2. This double comparison allowed us to set apart phenotypes caused by aberrant development from those due to developmental delay.

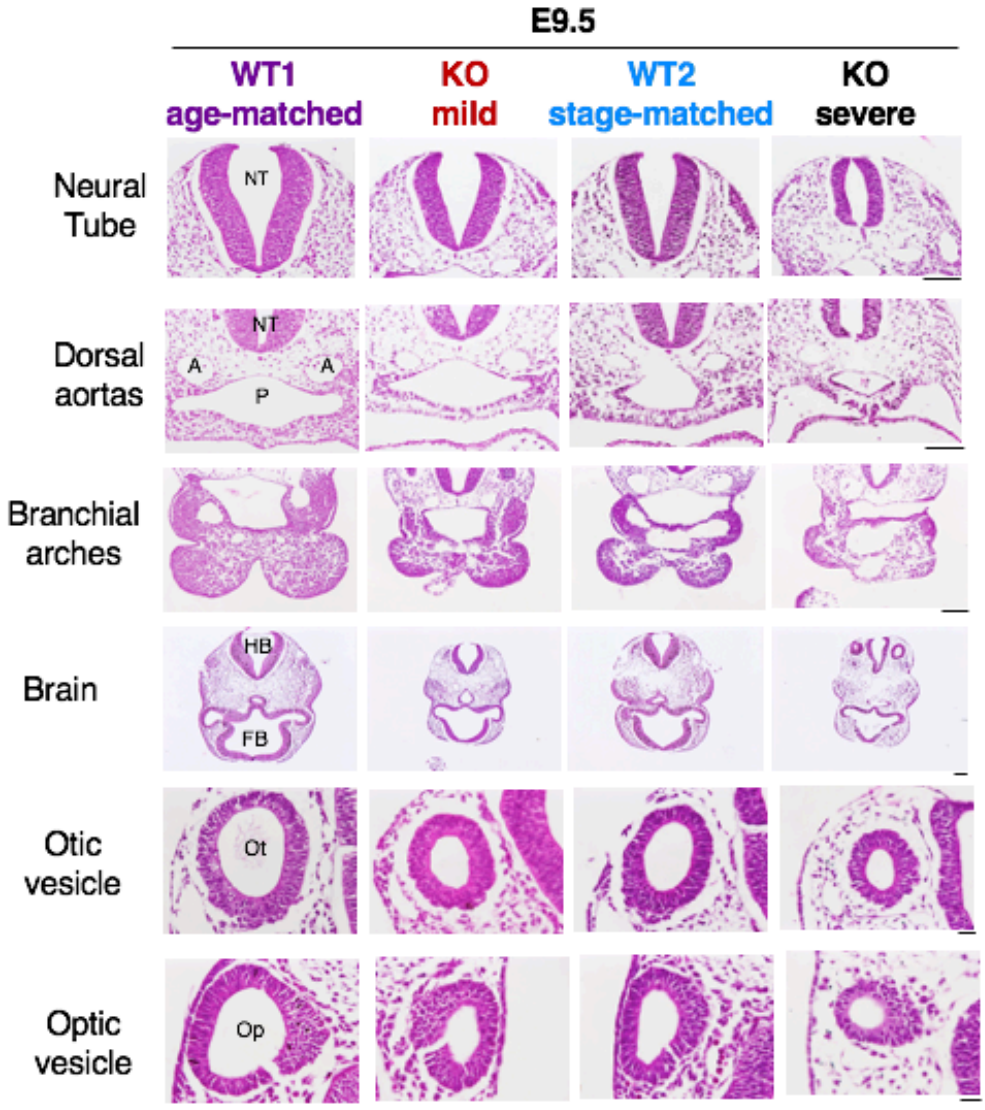


Figure R26. Global developmental defects in SA2-null embryos at E9.5.
Hematoxylin-eosin (H-E) stained transverse sections of male embryos of the indicated genotypes extracted at E9.5. Structures indicated on images are: neural tube (NT), aortas (A), pharynx (P), hindbrain (HB), forebrain (FB), otic vesicle (Ot) and optic vesicle (Op). Neural tube and dorsal aortas are taken from sections adjacent to the heart. Scale bars (valid for entire row): 100 μm for dorsal aortas, branchial arches and brain; 25 μm for otic and optic vesicles.

At E9.5, mutant embryos with the severe phenotype showed an aberrant morphology of all analysed structures (Fig. R26 and Fig. R27B). In contrast, most tissues and organs from SA2KO embryos with a milder phenotype did not show obvious malformations but they were clearly more similar to stage-matched (WT2) than to age-matched (WT1) controls (Fig. R26). Thus, the difference between SA2-proficient and -deficient littermate embryos appeared to be the result of developmental delay in the latter. A remarkable exception to this general trend was a selective defect in the developing heart.

4.6 Heart morphogenesis in the developing embryo requires cohesin-SA2

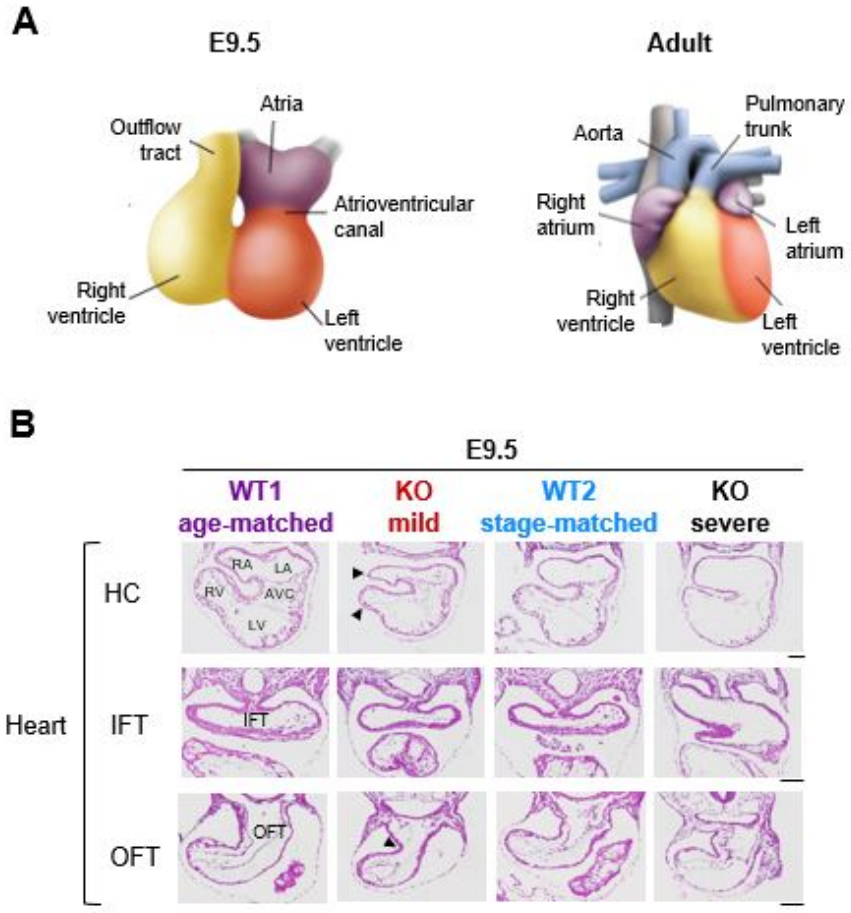


Figure R27. SA2-deficient embryos display heart defects.

A. Scheme showing the different heart regions in a murine heart at E9.5 and adult stages. Adapted from Epstein *et al.* 2015.

B. H-E stained transverse sections of different heart regions are shown: heart chambers (HC), inflow tract (IFT) and outflow tract (OFT). Regions within HC are: right atrium (RA), left atrium (LA), atrioventricular canal (AVC), right ventricle (RV); left ventricle (LV). Black arrowheads on HC indicate the position of the prospective septum between right and left chambers. Asterisks highlight the small size of the RV relative to the rest of the heart. White arrowhead points at the OFT curve. Scale bars (valid for entire row), 100 μm

At E9.5 the murine heart already presents a multichambered conformation as a result of linear heart tube extension. Two prospective ventricles and two prospective atria can be distinguished, although there is still no septation between them. The outflow tract (OFT) is a continuance of the ventricle that allows blood to flow out of the heart and will corresponds to the aorta and pulmonary trunk in an adult heart (Fig. R27A) (Kelly *et al.* 2014).

In E9.5 SA2KO embryos of the mild phenotype, the prospective atria and right ventricle were reduced in size compared to both controls (HC, heart chambers, in Fig. R27B).In contrast, no clear differences were found in the left ventricle Morphological defects were also observed in the OFT: SA2KO embryos showed an aberrant rightwards turning of the OFT at the junction with the ventricular myocardium when compared to stage-matched controls (OFT in Fig. 27B). However, the inflow tract (IFT) appeared normal (IFT, Fig. R27B). The defects described above were exacerbated in mutants with a severe phenotype, septum which displayed distended atria and ventricles with no visible indication of a future between right and left chambers and abnormal right ventricle development (last column in Fig. R27B). In this case, both the OFT and the IFT were distended.

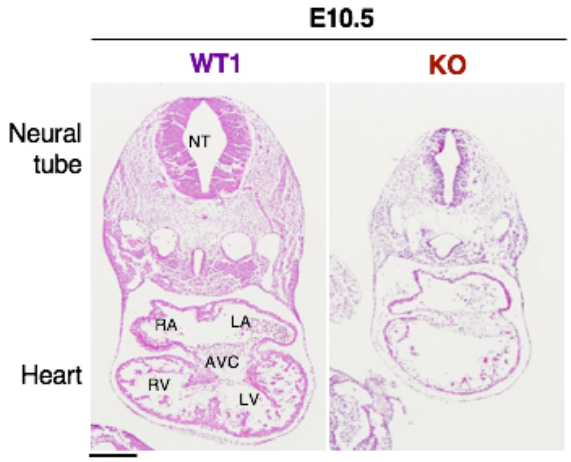


Figure R28. Severe cardiac anomalies in SA2 deficient embryos by E10.5.
H-E stained transverse sections of SA2KO and WT littermate embryos at E10.5. The sections shown encompass the neural tube and the heart chambers. Heart regions are indicated on the WT image as in A. Scale bar, 250 μm.

Despite the variable penetrance of the phenotype by E9.5, all mutants displayed severe cardiac anomalies by E10.5. The heart morphology resembled that identified in the severe phenotype at E9.5, with distended structures. Heart chambers lacked trabeculation and the atrioventricular canal was underdeveloped. Extensive necrosis and apoptosis were seen (Fig. R28).

In summary, SA2 mutant embryos display specific cardiac defects by E9.5 in addition to generalized developmental defects. We hypothesize that defective heart function may account for the embryonic lethality of SA2-null embryos. For further studies we proceeded only with SA2KO embryos of the mild phenotype to isolate the most primary defects.

4.7 SA1 and SA2 are ubiquitously expressed in E9.5 embryos

While SA1 and SA2 are expressed in all somatic cells, there are reports of different ratios of these somatic cohesin variants in different cell types (Kojic *et al.* 2018). We therefore asked whether the observed histological defects could arise from different expression patterns of SA1 and SA2 in embryonic tissues. For instance, cells expressing predominantly SA2 could be more sensitive to depletion of this cohesin variant.

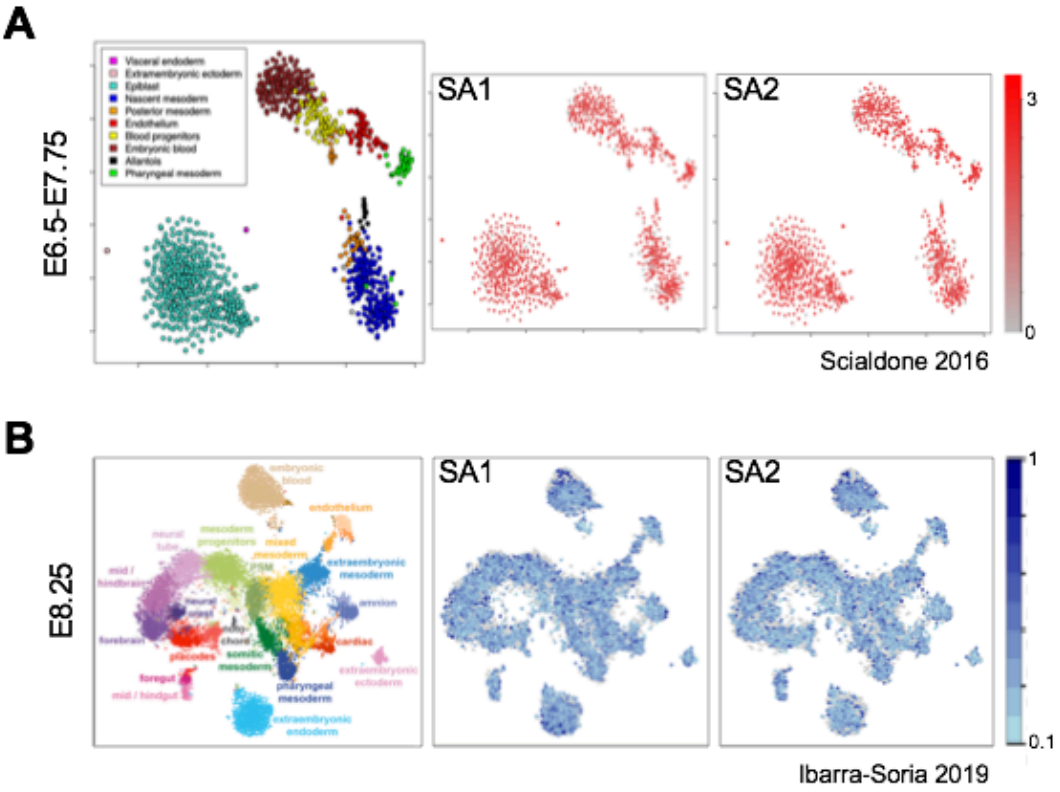


Figure R30. Similar expression patterns of SA1 and SA2 in early embryos.
T-distributed stochastic neighbor embedding (tSNE) plots from single-cell RNA-sequencing (scRNA-seq) data published in the indicated studies (left) and expression levels for SA1 and SA2 (middle and right). Different colors indicate different cell types.

A. Data from n=1,205 cells from in E6.5-E7.75 embryos identifies 10 cell clusters (Scialdone *et al.* 2016). Scale indicates $\log_{10}$ normalized counts for SA1 and SA2 transcripts. Data available at http://gastrulation.stemcells.cam.ac.uk/ scialdone2016.
B. Data from from n=19,396 cells from embryos at E8.25 stage identifies 33 cell clusters (Ibarra-Soria *et al.* 2018). Scale indicates $\log_{10}$ normalized counts +1 for SA1 and SA2 transcripts. Data available at http://marionilab.cruk.cam.ac.uk/ organogenesis.

We first looked into single cell RNA-sequencing (scRNA-seq) data from two published studies performed in mouse embryos at stage E6.5-7.75 (Scialdone *et al.* 2016) and E8.25 (Ibarra-Soria *et al.* 2018). The data are visualized using a dimensionality reduction technique called t-distributed neighbor embedding (tSNE; Van Der Maaten & Hinton 2008). For these datasets, the first two dimensions are shown and cells with similar transcriptional profiles are clustered into groups indicated by different colors. The relative expression levels of SA1 and SA2 were similar in all groups at E6.5-E7.75 (Fig. R30A) and E8.25 (Fig. R30B). Still unknown is how this RNA data corresponds with actual protein levels within the cell.

Since these data come from embryos at earlier stages, we also tested the relative levels of SA1 and SA2 in E9.5 embryos, in which cardiac defects were observed but prior to lethality. We checked protein levels by immunofluorescence in sections of WT E9.5 embryos with specific antibodies against SA1 and SA2.

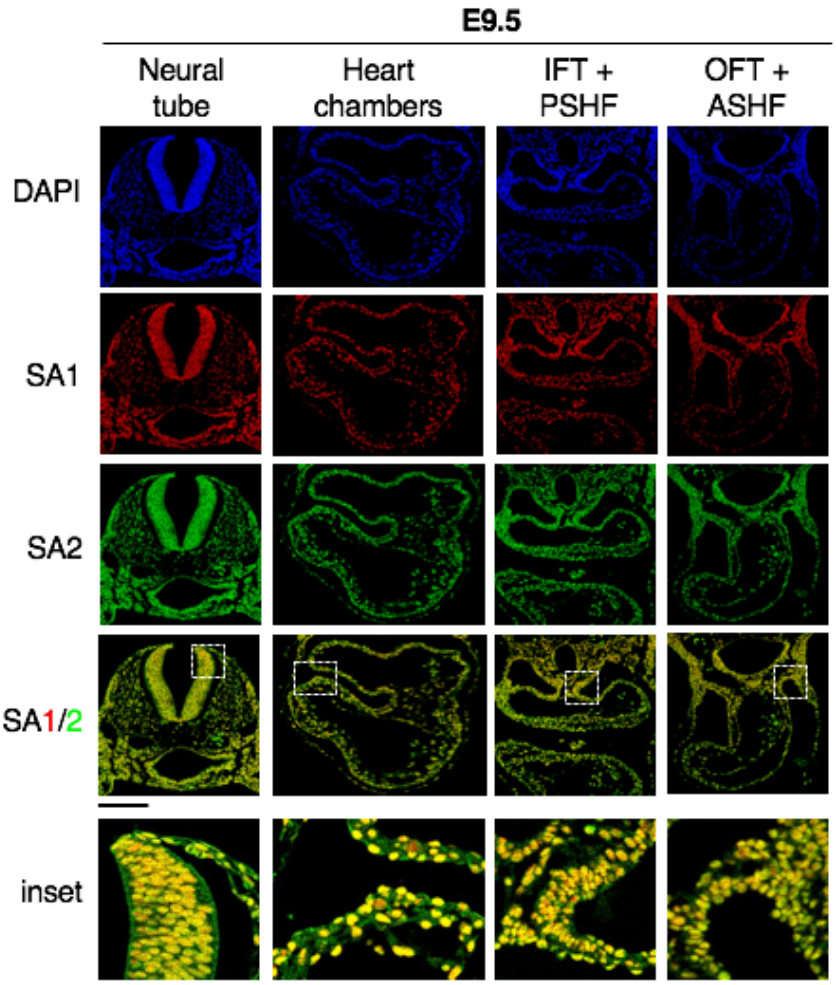


Figure R31. Distribution of cohesin variants in the E9.5 embryo.
Immunofluorescence co-staining of SA1 (red) and SA2 (green) in transverse sections containing the heart and neural tube of wild type E9.5 embryos. Nuclei are counterstained with DAPI (blue). Scale bar, 200 μm.

We detected widespread expression of SA1 and SA2 in all tissues (neural tube and heart are shown in Fig. R31). In the heart, we confirmed that both variants could be detected in all regions (heart chambers, IFT and OFT). Since we were using two distinct antibodies, one for SA1 and another for SA2, it was not possible to make conclusions about the absolute levels of the two proteins, but we could compare their ratio among different areas of the embryo. No obvious difference in the relative abundance of the two cohesin variants caught our attention when examining the heart and its adjacent regions. It is therefore unlikely that the specific cardiac defects observed at E9.5 are due to different expression levels of the two variants in heart and must instead result from a specific requirement of cohesin-SA2 in heart morphogenesis.

4.8 Decreased proliferation and increased apoptosis in *Stag2* mutant embryos

To explore the cellular mechanisms underlying the observed defects in heart morphogenesis, and given the defects in chromosome segregation and cell proliferation previously identified in SA2 KO MEFs, we went on to analyse proliferation and apoptosis in SA2 KO embryos by immunostaining. Transverse sections of E9.5 embryos of the mild phenotype as well as WT1 and WT2 embryos containing the different regions of the heart, or the neural tube for comparison, were selected and labeled with anti-phosphohistone H3 (H3P), Islet1 (ISL1), TUNEL and DAPI.

Histone H3 phosphorylated in Ser10 is abundant all over chromatin in mitosis and is also present in the heterochromatin of cells in late G2 phase, in which a dotted pattern is observed (Fig. 32A, left). Islet 1 is a transcription factor that identifies cardiac progenitors of the secondary heart field lineage and, together with morphological criteria, served to delimit the anterior and posterior secondary heart field regions (ASHF and PSHF, respectively) contributing to the heart. We developed a custom-made ImageJ macro that segmented the nuclei stained with DAPI and classified them into H3P-positive or H3P-negative (Fig. R32A, right). Examples of the stained sections are shown (Fig. R32B).

Figure R32. Analysis of cell proliferation in E9.5 embryos. >

A. Example of image processing and analysis using a custom-made ImageJ macro. The original (left) and processed (right) images are shown. An H3P signal above threshold is converted to a binary signal (red), and nuclei negative or positive for H3P are contoured in yellow or light blue, respectively.
B. Representative images of H3P staining in the indicated regions of the heart and the neural tube for comparison. Both original and processed images are shown, with corresponding insets. ISL1 (white) is only shown for regions that were defined based on its expression (ASHF and PSHF, marked with *). Scale bars (valid for entire column except insets), 100 µm. Quantifications are shown in Fig. R33.

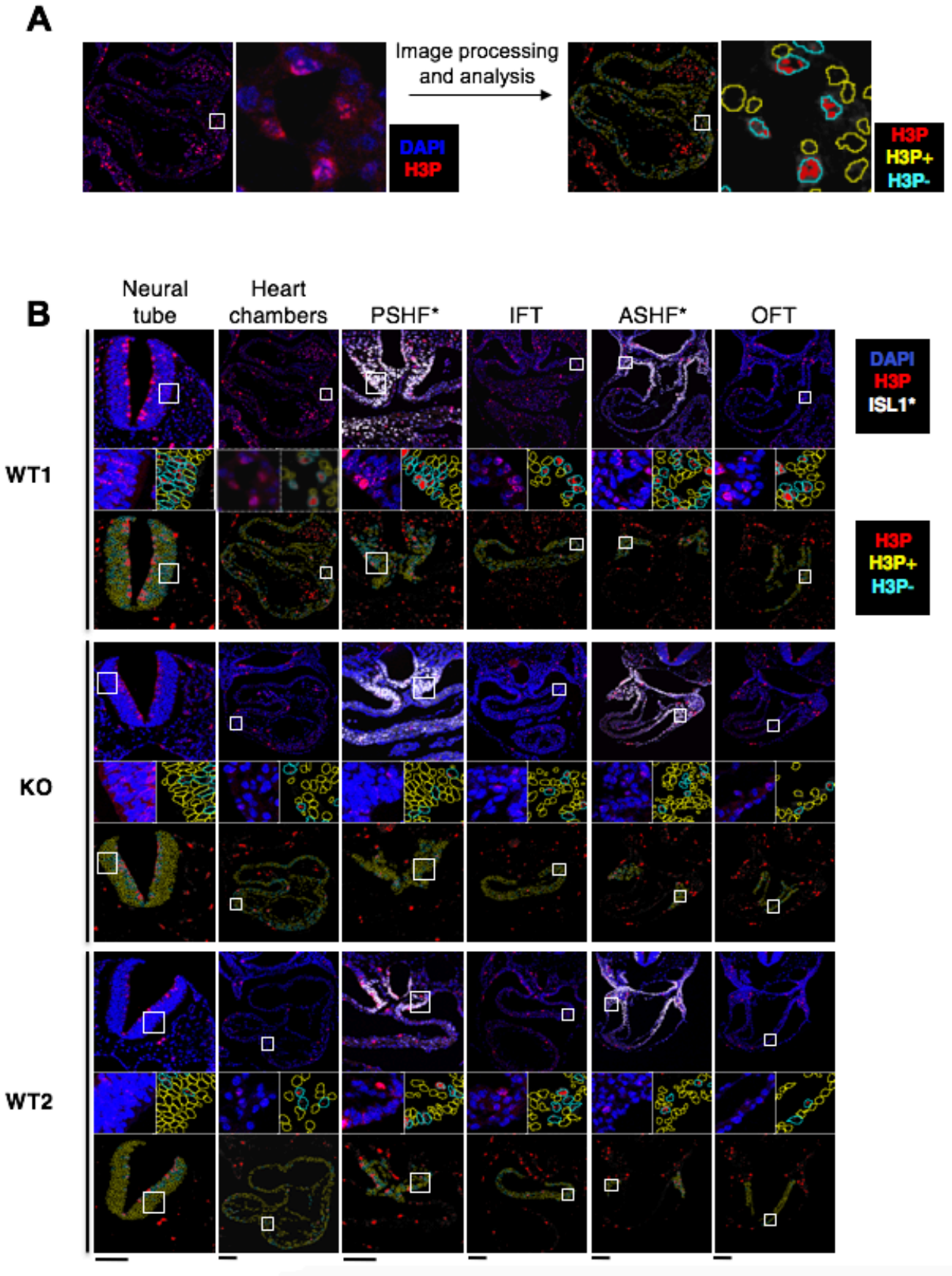


Figure R32. Analysis of cell proliferation in E9.5 embryos. Legend on previous page.

For each region, 9-12 sections coming from 3-4 embryos of each genotype were analysed and the values for each section were plotted (Fig. R33). The fraction of H3P-positive cells in the heart chambers (HC in Fig. R33) was significantly lower in the mutants as compared to their littermate controls (WT1), but similar to their stage-matched controls (WT2). The same was true for the anterior secondary heart field (ASHF) and the outfow tract (OFT), as well as the neural tube (NT), while in posterior secondary heart field (PSHF) and the inflow tract (IFT; Fig. R33).

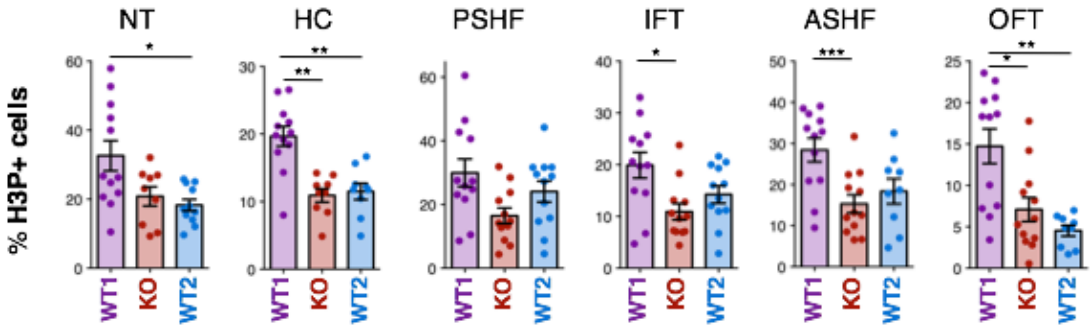


Figure R33. Reduced cell proliferation in SA2-null embryos.
Quantification of H3P-positive cells as readout for proliferation in WT1 (age-matched control), KO (mild phenotype) and WT2 (stage-matched control) E9.5 embryos in the indicated regions (as in Fig. 32). From 9-12 non-consecutive sections from 3-4 embryos were analyzed per genotype and region. Individual dots represent values for one section. Bars represent mean ± SEM. Kruskal-Wallis test and Dunn's multiple comparison post-test; *** P<0.001, **<0.01, * P<0.05, ns P≥0.05.

These same sections were co-stained with Terminal deoxynucleotidyl transferase dUTP nick end labeling (TUNEL). This assay allows labeling of cells that accumulate fragmented DNA, massively present in apoptotic cells (Fig. R34A). Overall, the fraction of TUNEL-positive cells was very low in the embryos and there was high inter-individual variability. In SA2KO embryos, we found an overall increased percentage of apoptotic cells compared to both controls. This difference was most notable in the neural tube, heart chambers and ASHF (Fig. R34B).

In sight of these results, we propose that the global developmental delay observed in SA2 null embryos at E9.5 could be due to a decrease in proliferative capacity of mutant cells, with some contribution from increased apoptosis.

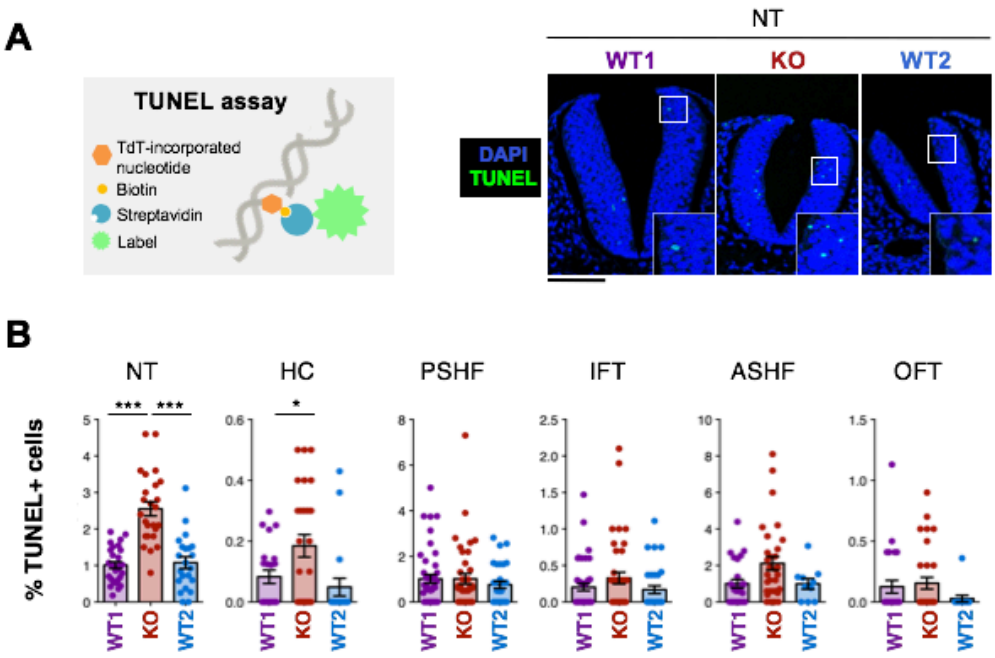


Figure R34. Increased apoptosis in SA2 null embryos.

A. Schematic representation of the TUNEL assay as a readout of apoptotic cells.
B. Representative images of TUNEL staining in neural tube of E9.5 embryos of the indicated genotypes.Scale bar, 100 μm.
C. Quantification of TUNEL-positive cells in sections as in Fig. 32-33. At least 10 sections from 3-4 embryos were analysed per genotype and region. Individual dots represent values for one section. Mean ± SEM are shown. Kruskal-Wallis test and Dunn's multiple comparison post-test; *** P<0.001, * P<0.05, ns P≥0.05.

4.9 Impaired deployment of progenitors into the heart tube in *Stag2* mutant embryos

While decreased proliferation might account for the global growth delay observed in the mutant embryos, it failed to explain why morphological defects were more evident in certain heart structures: the OFT, the right ventricle and the atria. Interestingly, these affected areas all derive from second heart field (SHF) progenitors, unlike the left ventricle, which derives from first heart field (FHF) (Kelly *et al.* 2014). The SHF is a population of cardiac progenitors of pharyngeal mesoderm origin that lies adjacent to the heart. By mid-gestation, addition of these progenitors to the linear heart tube drives its elongation and looping (Fig. R35A). The SHF progenitors populate the heart tube through the arterial and venous poles into the heart chambers: anterior SHF (ASHF) progenitors add onto the arterial pole through the OFT and posterior SHF (PSHF) progenitors migrate into the venous pole, through the IFT (Fig. R35B). As they migrate into the heart they differentiate into cardiomyocytes, which correlates with a progressively lower expression of the transcription factor Islet 1 (ISL1) (Francou *et al.* 2013; Kelly *et al.* 2014).

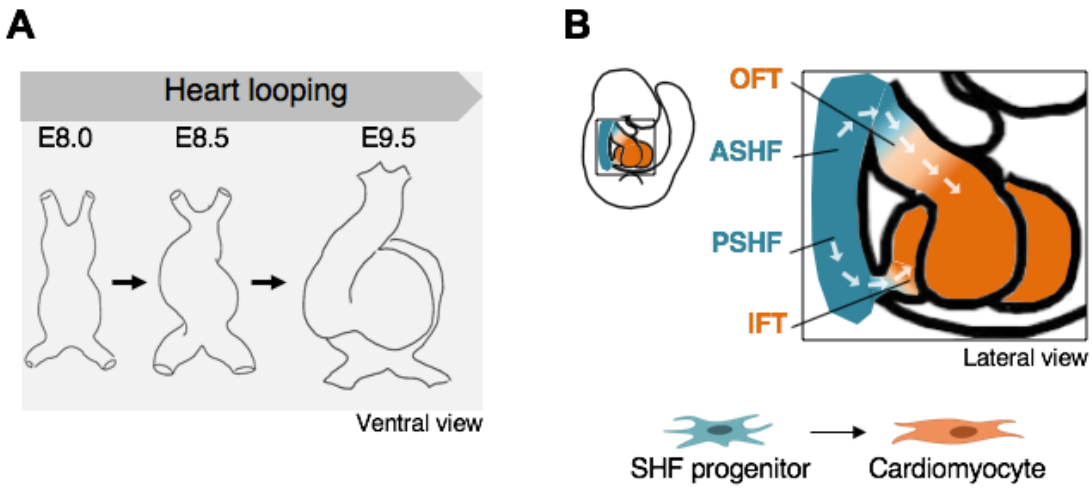


Figure R35. Two waves of progenitors in heart morphogenesis.

A. Schematic representation of heart looping. Between the stages E8 and E9.5, the linear heart tube progressively turns into a multichambered structure by asymmetric morphogenesis.

B. At E9.5, anterior secondary heart field (ASHF) and posterior secondary heart field (PSHF) progenitors migrate into the outflow (OFT) and inflow tracts (IFT) to populate the heart and contribute to its morphogenesis.

As part of the histological characterization, we measured the length of the OFT curves in equivalent sections of E9.5 embryos (Fig. R35A). In mutant embryos the length of the inner and outer curves was reduced compared to both controls (Fig. R35B). Together with the selectivity of heart defects, this led us to hypothesize a problem in the migration of ASHF progenitors into the OFT of SA2 mutant embryos.

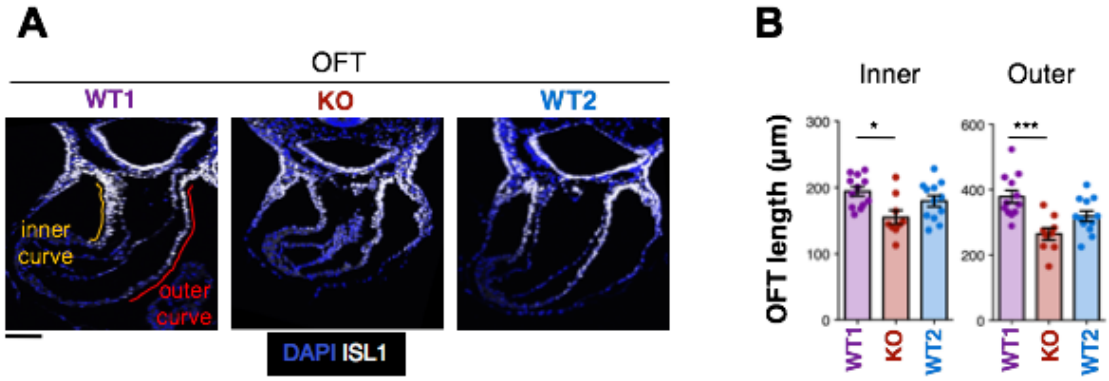


Figure R35. Reduced outflow tract length in SA2-null embryos.

A. Representative images of the outflow tract (OFT) in E9.5 embryos of the indicated genotypes. Inner and outer curve measurements are shown on the image for WT1 as an example.

B. Quantification of inner and outer OFT length in WT1, SA2KO and WT2 embryos at E9.5. 9-12 sections from 3-4 embryos were analyzed per genotype. Mean ± SEM are shown. Kruskal-Wallis test and Dunn's multiple comparison post-test; *** $P<0.001$, * $P<0.05$, ns $P\geq0.05$.

To test this possibility, we looked into the cellularity of the different heart regions. We quantified the total number of cells in the previously labeled E9.5 sections. In the neural tube, heart chambers and OFT, the cellularity of SA2KO embryos was significantly lower than in WT1 and similar to WT2 embryos (Fig. R36A), consistent with their reduced size and decreased proliferation rates (Fig. R26A and R33). In contrast, cell numbers in the ASHF were similar in SA2KO and WT1 littermates (Fig. R36A), despite mutants showing a reduced proliferation rate (Fig. R33). When we looked into the fraction of cells positive for ISL1, a transcription factor expressed in SHF progenitors that is progressively switched off as they migrate into the heart (Cai *et al*, 2013), we found that this fraction was similarly high in ASHF for all embryos. However, it decreased in the OFT of KO embryos compared to both controls (Fig. R36B), consistent with a failure of SHF progenitors to migrate efficiently into the OFT in SA2 null embryos.

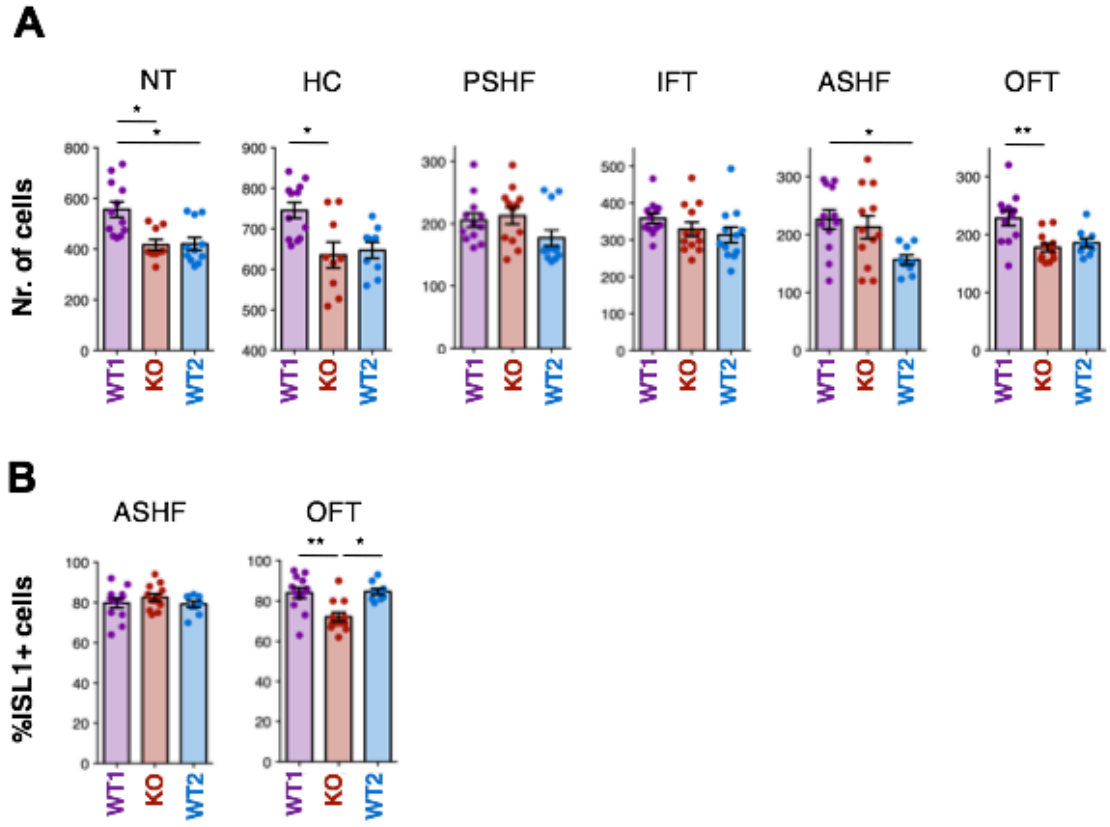


Figure R36. Evidence for impaired migration of anterior secondary heart field progenitors in the developing heart of SA2-null embryos.

A. Quantification of total number of cells in E9.5 WT1 (age-matched control), KO (mild phenotype) and WT2 (stage-matched control) embryos in the indicated regions: neural tube (NT), heart chambers (HC), posterior secondary heart field (PSHF), inflow tract (IFT), anterior secondary heart field (ASHF) and outflow tract (OFT).

B. Quantification of ISL1-positive cells in the ASHF and OFT regions of E9.5 WT1, KO and WT2 embryos. Mean ± SEM are shown. Kruskal-Wallis test and Dunn's multiple comparison post-test; ** P<0.01, * P<0.05, ns P≥0.05.

We performed a similar analysis in E8.5-E875 embryos (Fig. R37A), at the onset of SHF migration into the heart tube (Kelly *et al.* 2014) and focused on the OFT and ASHF. WT and SA2KO littermates with similar number of somites were longitudinally sectioned and labeled with an anti-ISL1 antibody (Fig. R37B). In terms of cellularity, there was a slight but non-significant reduction in the total number of cells in OFT and ASHF in KO embryos (Fig. R37C), despite no significant developmental delay at this stage (Fig. R26). The number of ISL1+ cells followed the same trend as at E9.5, with a decrease in the progenitors in OFT (Fig. R37D). Although the difference was not statistically significant, it is likely that the problem arises when the migration process starts.

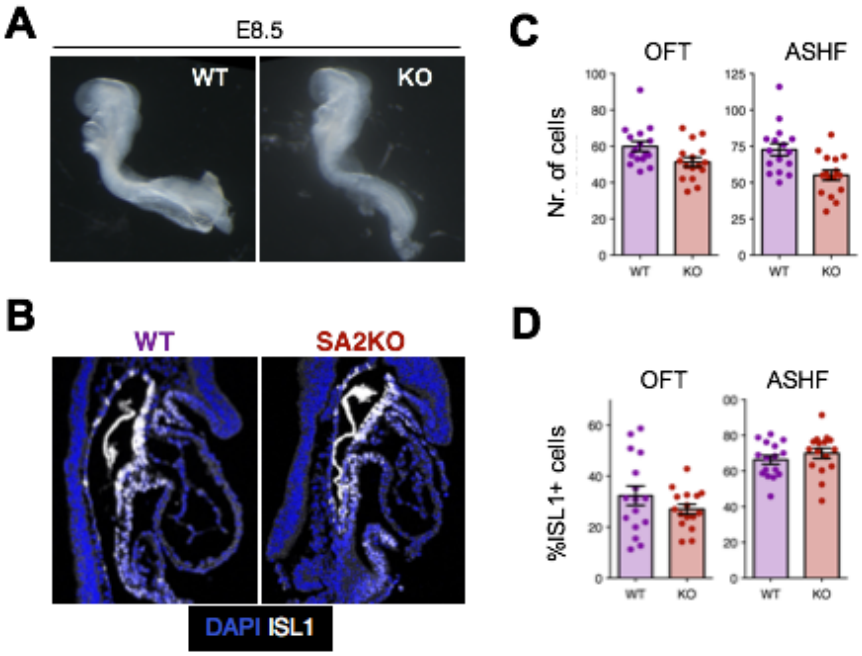


Figure R37. Migration of ASHF progenitors in SA2-null embryos at E8.5.

A. Representative images of whole-mount WT and SA2KO littermate embryos in embryos at E8.5.
B. Representative longitudinal sections containing the outflow tract (OFT) and anterior secondary heart field (ASHF) regions in WT and SA2KO embryos labeled with anti-ISL1 (white). Nuclei are counterstained with DAPI (blue).
C. Quantification of the cellularity in OFT and ASHF of E8.5 embryos. 12 sections from 4 embryos per genotype were quantified.
D. Quantification of ISL1-positive cells in OFT and ASHF of E8.5 embryos.

Altogether, these findings suggest that cohesin-SA2 loss results in accumulation of progenitors in the ASHF that fail to migrate efficiently into the heart tube, leading to morphological defects in ASHF derivatives such as the right ventricle and the OFT. Defects in migration of progenitors have already been suggested as the cause of heart defects in murine embryos and zebrafish deficient for the cohesin loader NIPBL (Muto *et al.* 2011; Santos *et al.* 2016).

4.10 Altered transcription of cardiac development regulators in *Stag2* mutant embryos

To address if the role of cohesin in gene regulation could contribute to the phenotypes described above, we compared the heart transcriptomes of E9.5 WT and SA2KO embryos by RNA-seq. To exclude variation related to developmental stage, we selected WT and SA2KO littermate embryos with similar number of somites. To identify tissue-specific changes, we compared heart and neural tube tissue adjacent to the heart. We used nine embryos of each genotype (3 pools of material from 3 embryos) to extract RNA.

First of all, we compared the transcriptomes of heart and neural tube in WT embryos. We identified 1,881 differentially expressed genes (DEGs; FDR<0.05) that allowed us to define a neural-enriched and a cardiac-enriched gene set for embryos at this stage (765 and 1,116 genes respectively; Fig. R38A). Gene Ontology (GO) analysis confirmed the functional specificity of these gene sets, as they were strongly enriched in pathways of tissue-specific processes (Fig. R38B).

When comparing the expression of these genes between WT and SA2KO embryos, there was a greater overall impact of SA2 loss on expression in the heart than in neural tube (heatmap in Fig. R38A). The transcriptional differences among the neural and cardiac gene sets in both tissues can also be represented in density plots reflecting the $\log_2$ fold change in SA2KO over WT. In these plots, it is clear that there is a larger deviation of both neural and cardiac gene sets in the heart of SA2KO embryos than in the neural tube. In addition, tissue-specific genes (e.g. cardiac genes in heart) were downregulated, while non-tissue-specific genes (e.g. neural genes in heart) were overall upregulated (Fig. R38C).

Pairwise comparisons between WT and SA2KO samples for each tissue identified 846 DEGs in heart but only 5 in neural tube (FDR<0.05; Fig. R39A), consistent with a larger deregulation in heart upon SA2 loss, as described above. Among the heart DEGs there were 222 and 112 genes from the cardiac and neural gene sets, respectively, and according to Fisher's exact test they were statistically overrepresented. This indicates that tissue-specific genes were preferentially affected by SA2 loss (Fig. R39B). Moreover, most cardiac genes were downregulated in the heart of SA2KO embryos, whereas the neural genes were upregulated therein (Fig. R39C).

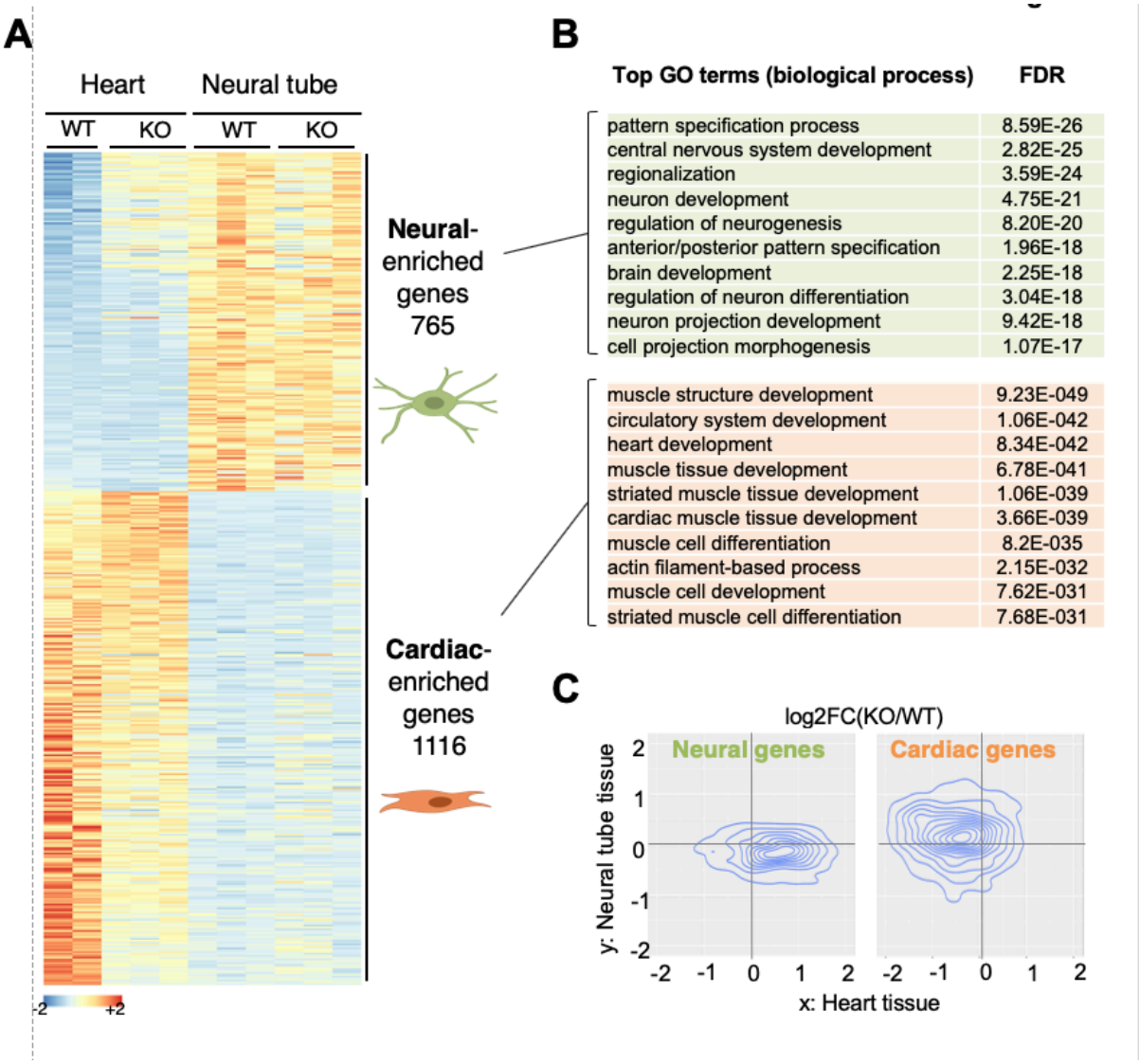


Figure R38. Tissue-specific transcriptional deregulation in SA2 null embryos.

A. Heatmap showing relative expression of 765 neural- and 1,116 cardiac- enriched genes in all RNA-seq samples. Gene sets were defined by differential expression between heart and neural tube samples (FDR<0.05).

B. Gene Ontology (GO) analysis of neural- and cardiac-enriched genes (shaded in green and orange, respectively). Top 10 significantly enriched pathways are shown ranked according to FDR (FDR<0.05).

C. Density plots displaying the differential expression (log_2 fold change; log_2FC) of gene sets from A in the two tissues.

These findings agree with the proposed role of cohesin-SA2 in tissue-specific transcription, promoting activation of genes specifying a tissue (i.e., cardiac genes in heart) and repression of alternative gene programs (e.g., neural genes in heart) (Kojic *et al.* 2018). This role is very clear in the heart, where loss of SA2 has a bigger impact, but can also be seen in neural tube to a lesser extent.

In addition, a closer look at the list of DEGs in heart revealed several cardiomyocyte markers and well-established regulators of SHF among the downregulated genes (Fig. R39A, right). For instance, Fgf8 and Hand2 contribute to the survival of ASHF progenitors while Wnt5a activity is critical for their deployment into the OFT. These roles are consistent with the defects described in the previous section.

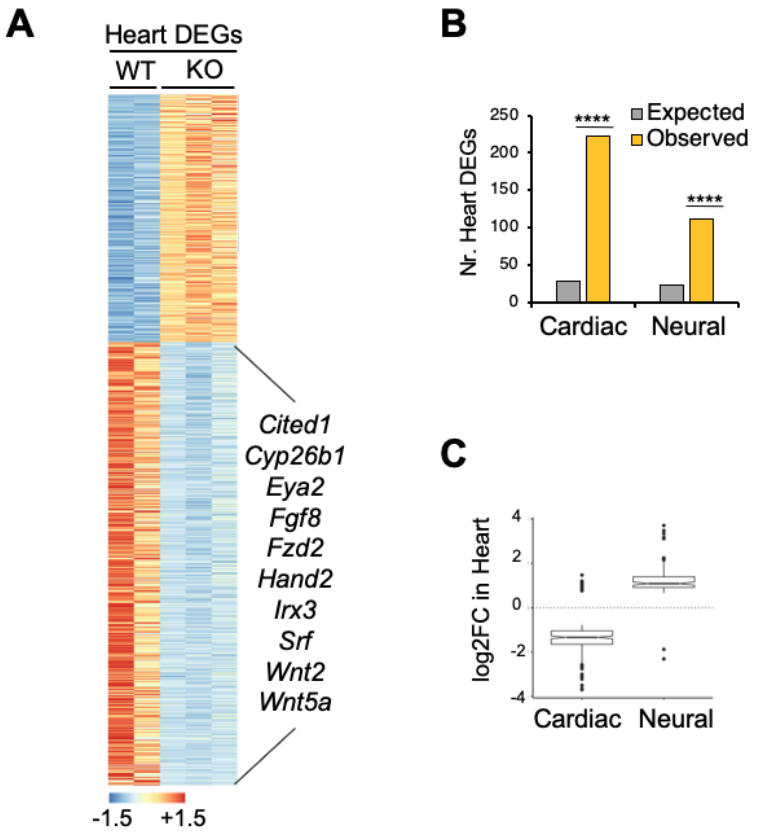


Figure R39. Differential expression in the heart of SA2 null embryos.

A. Heatmap of 846 DEGs between WT and SA2KO heart samples (FDR<0.05) from E9.5 embryos. Among the downregulated genes, we highlight some with established roles in cardiomyocyte differentiation and SHF regulation.
B. Expected versus observed number of cardiac and neural genes found among the heart DEGs. The total number of expressed genes is 21,653. Fisher's exact test; ****<0.0001 (p<2E-12).
C. Box plot of expression changes in the cardiac and neural genes identified as DEGs in heart

Taken together, our data suggest an important role of SA2 in the control of the early cardiac transcriptional programs that is not assumed by SA1 in the SA2KO embryos and which, together with decreased proliferation, contributes to the observed defects in heart morphogenesis.

Discussion

Discussion

As a major chromatin organizer, cohesin has a plethora of functions in DNA metabolism, including DNA replication, transcription, chromosome segregation and repair. A major challenge in cohesin biology has always been to understand how one complex can perform such varied functions. Actually, different isoforms exist of some cohesin subunits and accessory proteins, giving rise to different combinations that can potentially divide the labor. Diversifying composition is a common way for cells to fine-tune functions of complexes, and a close example can be found within the SMC family: condensin. Vertebrate cells contain two versions of condensin, I and II, that participate in chromosome condensation in a stepwise manner (Hirano 2005).

In this Thesis we have focused on cohesin-SA1 and cohesin-SA2. SA1 and SA2 share a high degree of sequence homology and their specificity has been largely neglected. Previous studies have addressed the differential role of these two cohesin variants in basic cohesin functions (Canudas & Smith 2009; Remeseiro *et al.* 2012a), but more recently studies by our group and others have uncovered unique roles beyond cohesion (Cuadrado *et al.* 2019; Kojic *et al.* 2018; Viny *et al.* 2019). We have addressed unique roles of cohesin-SA2 *in vitro* as well as during embryonic development to further understand functional specificities.

1. SA1 allows survival of SA2-null cells *in vitro* with mild proliferation defects

While entailing some phenotypes, our studies in MEFs show that depletion of SA2 does not compromise cell survival. Consistent with previous studies in other cellular systems (Benedetti *et al.* 2017; Liu *et al.* 2018; van der Lelij *et al.* 2017), this confirms that the presence of either cohesin variant is sufficient to sustain cell viability *in vitro* and that SA1 and SA2 may compensate each other for the essential functions of cohesin. In this regard, the phenotype of SA1 or SA2 depletion may be due to reduced cohesin rather than specific defects. Only a fraction of the total amount of cohesin present in the cell is sufficient to promote cohesion during mitosis, as suggested by studies in budding yeast and Drosophila cells showing that cohesin levels need to drop below 13% or 20%, respectively, for cohesion loss to manifest (Carvalhal *et al.* 2018; Heidinger-Pauli *et al.* 2010). We often observe upregulation of SA1 protein levels in MEFs after SA2 depletion, suggesting that the cell tries to compensate total cohesin levels. Therefore, there might

be some specificity in the mechanisms promoting cohesion, despite the overall functional redundancy.

As far as the basic functions of cohesin are concerned, SA2-null cells show phenotypes that are not unlike those in SA1-null cells (Remeseiro *et al.* 2012a). These include a reduction in proliferation, to which multiple mechanisms could contribute. Transcriptional analysis show that “Cell cycle”, “DNA replication” and multiple DNA repair pathways were all downregulated in SA2-depleted MEFs.

Our analyses suggest that SA2 deficient MEFs progress slower through the cell cycle and show a mild defect in S phase progression. These results are in stark contrast with a recently published study that claims a strong role of SA2 in stability and progression of the replication fork in non-transformed cells. SA2 ablation in RPE cells caused intra-S-phase arrest, replication fork collapse and accumulation of double strand breaks (DSB), leading to a senescent phenotype (Mondal *et al.* 2019). This same study showed how SA2-mutant cancer cell lines were more sensitive to DNA damaging chemotherapeutic agents and inhibitors of DSB-repair, compared to SA2-proficient ones. This is consistent with earlier reports of increased sensitivity of cohesin to PARP inhibitors(Bailey *et al.* 2014; Mclellan *et al.* 2012; O’Neil *et al.* 2013). However, no strong S-phase defect was observed in our cellular model upon depletion of SA2. Differences in total and relative amounts of cohesin variants found in different cell lines could explain this, highlighting the importance of the system used.

In addition, we observed defects in centromeric cohesion, agreeing with the previously described role for cohesin-SA2 in HeLa and mouse C2C12 cells (Canudas & Smith 2009; Remeseiro *et al.* 2012a). The defects were mild in nature, with cells usually showing increased distance between sister centromeres as opposed to fully separated ones, that could translate into reduced resistance to spindle forces. The presence of SA1 and/or alternative mechanisms contributing to cohesion likely make up for the loss of SA2. A fraction of SA2-deficient cells exhibits segregation defects and aneuploidies, which could lead to inviable cells and contribute to reduced proliferation.

Despite reasonable maintenance of cohesion and overall viability in unchallenged conditions, SA2 deficient MEFs may be more sensitive to DNA damage since previous studies have shown a preferential role of cohesin-SA2 over cohesin-SA1 in DNA repair (Kong *et al.* 2013; Meisenberg *et al.* 2019). This possibility remains to be addressed.

2. The different genomic distribution of cohesin-SA1 and cohesin-SA2 reveals two classes of cohesin positions

While both SA1 and SA2 may compensate for the essential functions of cohesin, at least *in vitro*, this might not be the case for other non-essential functions. As a major determinant of cohesin function in genome organization, we studied the genome-wide distribution of cohesin-SA1 and cohesin-SA2 in MEFs. We identified two major types of cohesin positions with differences in their colocalization with CTCF and their enrichment in functional genomic elements: common cohesin positions and SA2-only positions (illustrated in Fig. D1).

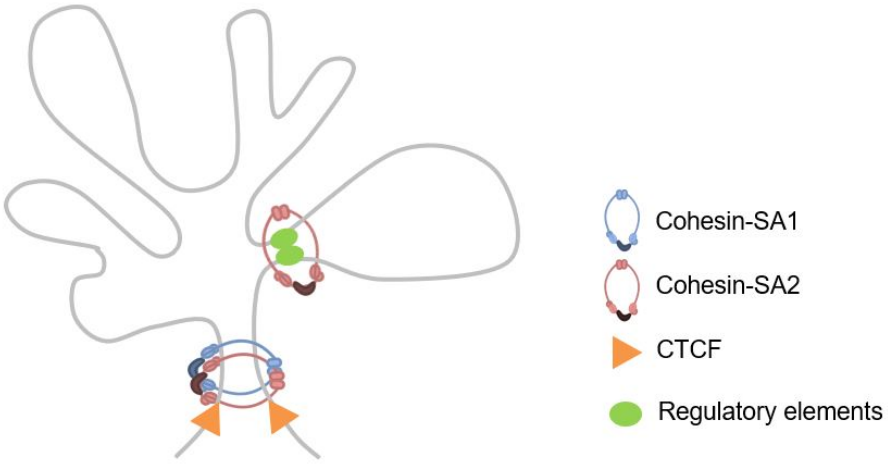


Figure D1. Common and SA2-only cohesin positions.
Simplified illustration of the two types of cohesin positions identified in MEFs. Common cohesin positions are shared by cohesin-SA1 and cohesin-SA2 and colocalize with architectural protein CTCF, forming large structural loops corresponding to TADs or sub-TADs. SA2-only positions exist independently of CTCF creating more local contacts that are highly enriched in regulatory elements.

Both cohesin-SA1 and cohesin-SA2 are present at common positions, which overlap with CTCF and largely correspond to the “insulator” chromatin state. At least a subset of these positions likely demarcate TADs or subTADs. The presence of CTCF is the major determinant of cohesin localization at these sites, as removing CTCF drastically decreases cohesin occupancy at these sites. They likely pose an obstacle for cohesin progression by loop extrusion (Busslinger *et al.* 2017; Haarhuis *et al.* 2017). The common nature of these position suggests that cohesin-SA1 and cohesin-SA2 can be similarly stopped here, and do so independently of the other cohesin variant, as was confirmed in SA1- and SA2-deficient MEFs.

In contrast, SA2-only positions lack SA1, are exclusively made up of cohesin-SA2 and do not overlap with CTCF. They are highly enriched in chromatin states corresponding

to active enhancers and promoters. CTCF-independent cohesin positions had been previously described and associated to binding of tissue-specific transcription factors, yet they were not clearly attributed to a specific cohesin variant (Faure *et al.* 2012; Kagey *et al.* 2010; Schmidt *et al.* 2010). Work from our group has shown that cohesin-SA2 is the variant present at non-CTCF cohesin sites in multiple cell types (this Thesis, Cuadrado *et al.* 2019; Kojic *et al.* 2018).

Consistent with their different distribution, depletion of either isoform has a different impact on gene expression that affects a wide array of cellular processes, as shown by our transcriptome analysis. In fact, depletion of SA1 and SA2 in MEFs has strikingly opposite effects in terms of deregulated gene sets. Interestingly, a similar trend was observed in mESCs (Cuadrado *et al.* 2019). These results point to distinct mechanisms of gene regulation by cohesin-SA1 and cohesin-SA2, which remain to be elucidated. It also remains unclear which transcriptional effects are a consequence of global alteration of TAD structure, which result from alteration of a local contact involving the gene anf its regulatory element requiring cohesin and which correspond to secondary effects.

The two types of cohesin positions identified exist among multiple cell types, including human cell lines of different embryonic origin, mESCs and mouse hematopoietic stem cells (Cuadrado *et al.* 2019; Kojic *et al.* 2018; Viny *et al.* 2019). This highlights their importance and points to a universal mechanism of genome organization by cohesin variants. What remains enigmatic is how different cohesin complexes can occupy different positions.

3. Distinct dynamic behavior of cohesin-SA1 and cohesin-SA2 and its functional relevance

Common cohesin positions have a high cohesin occupancy and correspond to narrow peaks. SA2-only positions, in contrast, have lower cohesin occupancy and display broader profiles. Since ChIP-seq is a population study, it shows the cumulative binding across many cells at the time of crosslinking. Therefore, robust cohesin positions likely reflect the presence of cohesin at these sites in a large number of cells from the population. These cells could present either cohesin-SA1 or cohesin-SA2 at these sites, or both variants at the same time. Re-ChIP experiments have shown that at least in some common positions more than one cohesin complex can be present (Kojic *et al.* 2018). Weaker cohesin positions, on the other hand, could correspond to sites found in a

minority of cells at a given time. It is likely that these positions are more transient and therefore overall less represented in the population.

The weaker nature of SA2-only positions suggested to us that cohesin-SA2 present at SA2-only positions was more dynamic. Biochemically, we showed that overall SA1 was less sensitive to salt extraction than SA2, which was more readily destabilized. This experiment allowed us to conclude that SA2 is less tightly associated to chromatin but did not speak directly to cohesin dynamics. iFRAP experiments allowed us to address this directly and demonstrated that cohesin-SA2 has a more dynamic behavior than cohesin-SA1.

We suspected that the molecular reason behind this could be related to different susceptibilities to cohesin release factors such as WAPL. We observed a decreased interaction of cohesin with WAPL upon depletion of SA2, suggesting that this factor preferentially associates to cohesin-SA2. Immunoprecipitation experiments in human cell extracts have shown a similar preferential interaction of SA2 with WAPL (Kojic *et al.* 2018). Moreover, upon WAPL depletion SA2-only positions are selectively reduced, suggesting that they rely on cohesin dynamics.

Interestingly, a highly related study in our lab has shown an increased WAPL/SA2 ratio associated to SA2-only positions compared to common positions, by ChIP-qPCR in MCF10A cells (Kojic *et al.* 2018). These data indicated that the higher susceptibility to WAPL of cohesin-SA2 was not general to the entire pool of cohesin-SA2, but rather restricted to the population found at SA2-only sites. This raised the possibility that cohesin at common positions, comprising cohesin-SA1 and/or cohesin-SA2, could be protected from WAPL unloading. These positions are sites of CTCF binding. Intriguingly, the region of SA2 that interacts with CTCF has been narrowed down to residues 162-290 by in vitro pull-down assays (Xiao *et al.* 2011), and residues very close to this region (positions 290, 326 and 330) are key for binding to WAPL (Hara *et al.* 2014). This raises the possibility of competitive binding of CTCF and WAPL to the SA subunit of cohesin. Although the region in question is highly conserved between SA1 and SA2, the preference to either factor could be slightly different for both variants, with SA1 preferring CTCF and SA2 preferring WAPL.

With these data in mind, we envision a model in which cohesin extrudes chromatin until released by WAPL or until it reaches CTCF sites, where WAPL accessibility is impaired, possibly by WAPL-CTCF competition. We propose that cohesin-SA1 can reach CTCF

boundaries with a higher frequency than cohesin-SA2 and is therefore stabilized on chromatin for longer times. Cohesin-SA2, on the other hand, is more susceptible to WAPL release during loop formation, leading to a higher turnover on chromatin. However, once it reaches CTCF sites, interaction with CTCF equally stabilizes it by protection from WAPL (Fig. D2). It is also possible that physical stacking of several cohesin rings at a single common position also physically hinders WAPL accessibility.

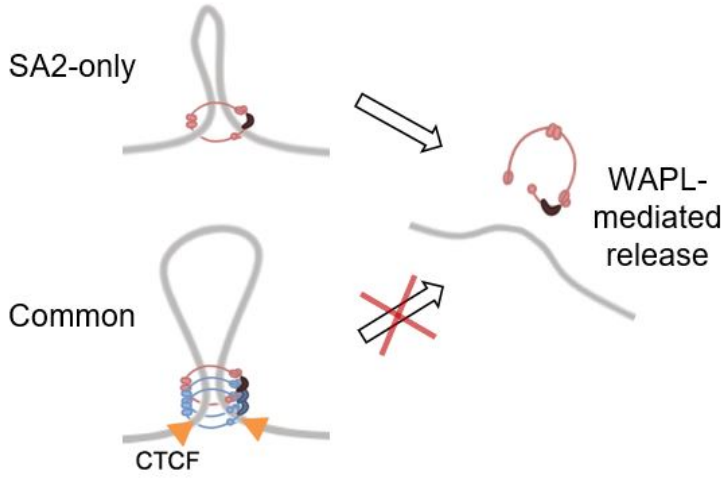


Figure D2. Cohesin-SA2 at SA2-only positions is more susceptible to WAPL-mediated release.
Model proposing that cohesin-SA2 is more sensitive to WAPL-mediated release from chromatin, leading to a higher turnover. The presence of CTCF or potential cohesin stacking at common positions could protect them from WAPL activity, leading to more stable positions.

We believe the distinct dynamic behavior and genomic localization to be directly relevant to the role of cohesin in transcription. Establishment of different types of contacts can dictate gene expression and explain why depletion of SA1 and SA2 have different effects on transcription.

Considering the differences between cohesin at common and SA2-only sites in terms of association with WAPL and CTCF, we postulate that cohesin dynamics should change upon depletion of CTCF or cohesin removal factors WAPL (Fig. D3). In the absence of CTCF, cohesin would not be protected from WAPL-mediated release and cohesin-SA1 would adopt a more dynamic behavior. This would agree with our observation that cohesin-SA1 can occupy SA2-only positions when CTCF is absent. Upon depletion of WAPL it has already been shown that cohesin's residence time on chromatin is extremely increased (Tedeschi *et al.* 2013), and we think this would primarily affect SA2, just as it would in a PDS5 KO situation. A preprint showing how SA1 and SA2 dynamics change in these backgrounds has been recently posted (Wutz 2019), confirming some of these hypotheses.

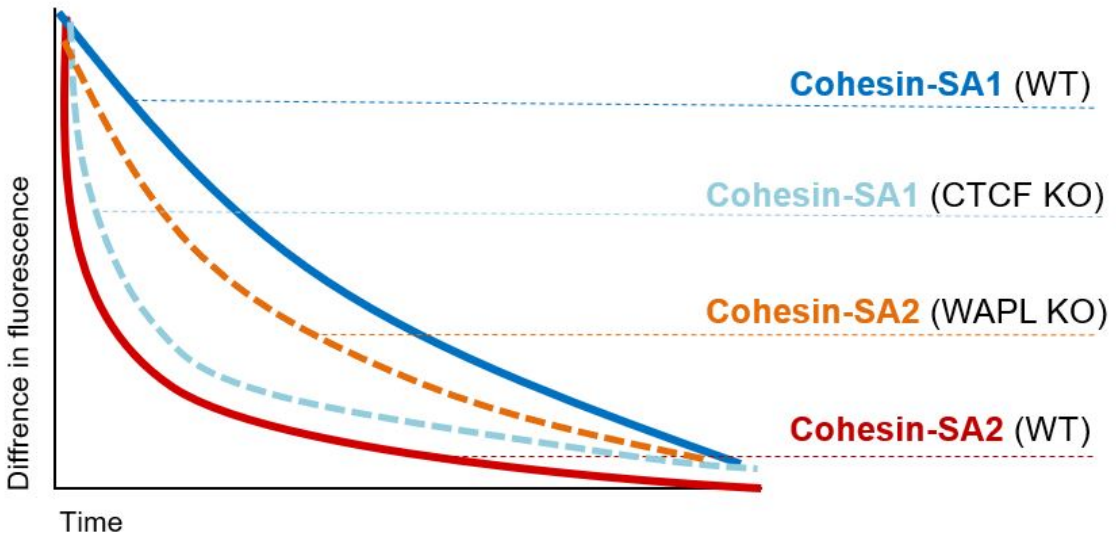


Figure D3. Potential behavior of SA1 and SA2 in iFRAP experiments in different backgrounds.
Blue and red continuous lines represent results obtained in our iFRAP analysis for SA1-GFP and SA2-GFP, respectively. CTCF KO (light blue dashed line) likely makes SA1 more similar to SA2 in terms of dynamics, as cohesin at common positions would no longer be protected from WAPL-mediated removal. In an opposite fashion, depletion of WAPL (orange dashed line) would make SA2 behave more like SA1.

In addition to differential dynamics, we must not forget the importance of specific interacting partners. Non-CTCF-cohesin, now identified as cohesin-SA2, has previously been shown to correlate extensively with binding of tissue-specific transcription factors by ChIP-seq (Faure *et al.* 2012; Schmidt *et al.* 2010). In our group we have also shown direct interaction of SA2 with transcriptional regulators like YAP1 and ZMYM2 in both HeLa and MCF10A cells (Kojic *et al.* 2018). Moreover, we have shown that SA2-only positions are enriched in binding motifs from specific transcription factors, such as the AP-1 and TEAD families, that are different to those found in common positions. Therefore, it is plausible that SA2 can uniquely interact with certain transcriptional regulators. The elucidation of the full structure of SA subunits, comprising the least conserved regions between SA1 and SA2, may give important clues in this regard.

4. Both cohesin variants are essential for mouse embryonic development

In contrast to the redundancy and functional compensation of SA proteins *in vitro*, mouse embryos require both proteins to fulfill their development. Previous work from our group showed that constitutive inactivation of *Stag1* in the germline is embryonic lethal, with incomplete penetrance, starting at stage E11.5. SA1-null embryos displayed a severe

developmental delay but no obvious organ malformation (Remeseiro *et al.* 2012b). In this Thesis we have shown that inactivation of *Stag2* in the germline leads to an earlier and more abrupt lethality, starting at E9.5. The phenotype is associated with a developmental delay and a broad, subtle tissue disorganization, with the exception of severely impaired heart development. In first instance, this tells us that successful development relies on the presence of both cohesin variants, and that neither can compensate for the other in its absence. Thus, during the formation of a multicellular organism, both SA1 and SA2 have unique and non-redundant roles.

Both SA1- and SA2-null embryos share a common phenotype: developmental delay. In both cases this could be inferred from embryo size but it was also precisely quantified for SA2-null embryos through somite counts. As in cell cultures, reduced proliferation was observed in SA2-null embryos, a defect that likely contributes to the growth retardation. This could also happen in SA1-null embryos, although it was not assessed. In both cases, there is a variable penetrance of the phenotype. A plausible explanation is the variable compensation of expression of the remaining cohesin variant, and as a result, different global cohesin levels between embryos. We have indeed seen variable upregulation of SA1 protein levels in MEFs upon SA2 depletion, suggesting the same variable compensation could happen *in vivo* and lead to different outcomes.

In SA2-null mice lethality occurs quite early and no embryos survive beyond E10.5. Other mouse models partially deficient for genes crucial in cell proliferation survive to later stages of development. For example, mice carrying hypomorph alleles of the MCM3 subunit of the replicative helicase survive until E16.5-18.5, and their associated lethality seemingly results from impaired expansion of hematopoietic precursors (Alvarez *et al.* 2015). Mutant mice in the centrosome component Cep57 that display more severe chromosome segregation anomalies than those reported for SA2-null MEFs also survive to birth (Aziz *et al.* 2018). Therefore, we believe it is unlikely that defects in proliferation alone account for the lethality of SA2 KO embryos.

5. Heart defects in SA2-null embryos

In addition to the systemic developmental delay, SA2-null embryos present a very overpowering defect in the developing heart. Defects were observed in other tissues, but these were coupled to the developmental delay, which was clarified by comparison with stage-matched controls. Proliferation defects were not specific to the heart and were

observed in all regions analyzed, suggesting that additional processes are altered during cardiac morphogenesis in the absence of SA2. Interestingly, histological defects did not affect the heart entirely but were restricted to the right ventricle, outflow tract and atria, regions that all derive from the second heart field (SHF).

At early stages of embryonic development, a linear heart tube is formed from first heart field (FHF) progenitors. At approximately stage E8.5, contribution of the second heart field (SHF) pool of progenitors to anterior and posterior poles of the heart induces heart looping, breaking the symmetry of the heart and giving rise to a multi-chambered structure. This event is crucial for heart morphogenesis and relies on rapid proliferation of SHF progenitors that migrate into the heart tube as they differentiate (Francou *et al.* 2013).

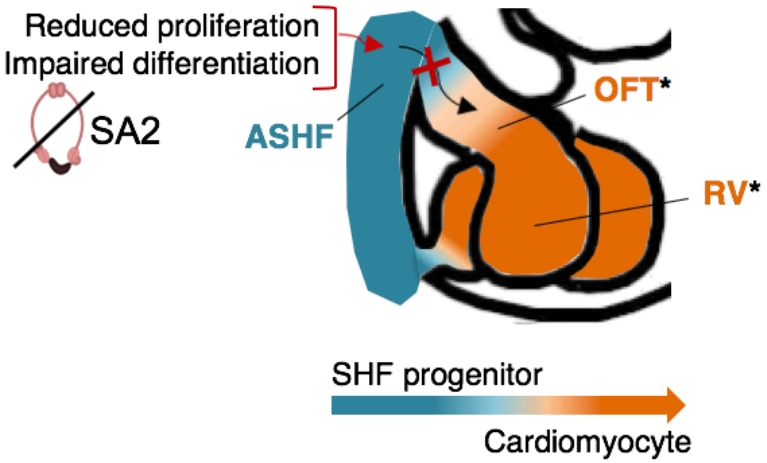


Figure D4. Impaired deployment of SHF progenitor into the heart tube of SA2-null embryos.
Impaired differentiation caused by absence of SA2, as well as decreased proliferation, likely cause the impaired migration of ASHF progenitors into the heart tube of SA2-mutant embryos, leading to morphological defects in regions directly derived from this population (marked with *).

Our detailed analysis of SHF cells and cellularity suggest that migration of anterior SHF (ASHF) progenitors is impaired in SA2-null embryos (Fig. D4). This could explain morphological defects in the regions that derive from these progenitors, for example the shortened OFT or the hypoplastic right ventricle. This idea was reinforced by the finding of transcriptional downregulation of key regulators of SHF in the heart of SA2-deficient embryos. To further confirm this hypothesis, lineage tracing experiments *in vivo* would be very suitable, using a SHF-specific enhancer-driven reporter system to directly monitor the contribution of SHF cells to the heart tube.

As discussed above, increasing evidence supports the notion that the presence of cohesin-SA2 at regulatory elements independently of CTCF promotes cell type-specific transcription, a function that cohesin-SA1 cannot compensate (Cuadrado *et al.* 2019; Kojic *et al.* 2018; Viny *et al.* 2019). From these studies, we know that common cohesin positions are highly conserved, while SA2-only sites are highly cell-type specific (Kojic *et al.* 2018). In mESCs it was shown that cohesin-SA1 contributes mainly to global chromatin architecture while cohesin-SA2 had additional roles in mediating contacts important for repression of lineage commitment genes and maintenance of pluripotency (Cuadrado *et al.* 2019). Altogether, this suggests that cohesin-SA2 plays a crucial role in enforcing the correct transcriptional programs in accordance to developmental state and cell type. We hypothesized that cohesin-SA2 would be highly relevant during morphogenesis to rewire transcription and promote lineage commitment. Comparison of transcriptomes of heart and neural tube from WT E9.5 embryos allowed us to define tissue-specific expression profiles corresponding to this stage of development. Consistent with this idea, transcriptome analyses uncovered altered tissue-specific transcription patterns in SA2-null embryonic hearts, with lower expression of cardiac genes and de-repression of genes from the neural lineage. This trend was also observed in neural tube, although this tissue presented an overall smaller deregulation (Fig. D5). In our experiment, however, the dissection of whole hearts may dilute defects in specific populations like the SHF. Thus, this study would greatly benefit from single-cell RNAseq.

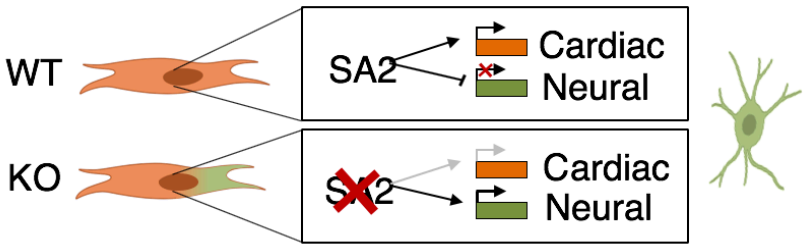


Figure D5. Loss of cell identity in the absence of SA2.
In normal conditions, SA2 is responsible for inducing tissue-specific transcription programs while repressing genes from other lineages. In the absence of SA2 there is a reduced lineage commitment and aberrant activation of genes from other lineages.

Overall, we propose that defects in both proliferation and lineage specification driven by cohesin-SA2 contribute to the heart abnormalities observed in the SA2-deficient embryos. This was previously suggested in NIPBL-deficient mouse and zebrafish embryos (Kawauchi *et al.* 2016) or zebrafish with reduced cohesin levels that often failed to loop (Santos *et al.* 2016; Schuster *et al.* 2015).

However, it is still puzzling how aberrant development was so specific to the heart. We believe that heart defects are the main cause of death in SA2-null embryos. The heart is one of the first organs to start differentiating in the embryo, and the first one to become functional (Bruneau 2013). Heart failure would make embryos unable to sustain further development, thus masking potential defects in other organs that arise later in development. We predict that bypassing the heart defect, for example by restoring SA2 expression specifically in the heart, would set back the lethality and allow the observation of developmental defects elsewhere in the embryo. This is backed up by the fact that in the neural tube we also observed deregulation of tissue-specific genes, although the magnitude of the changes was not significant.

A possibility is that cohesin requirements change during development in different cell populations. During a particular developmental window in a particular cell type, the requirement for cohesin could peak, and cells in this window would be particularly sensitized to cohesin depletion (Fig. D6). This model was proposed by Schuster *et al.* 2015.

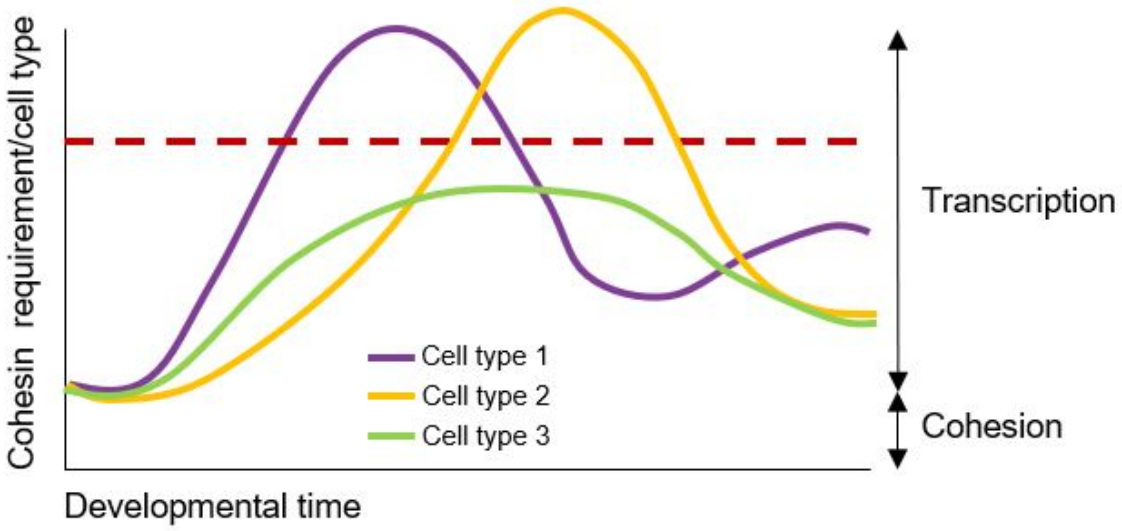


Figure D6. Model of fluctuation in cohesin requirements in different cell types during
Model illustrating how cohesin requirements (purple curve) can change during development in different cell types. A small fraction of cohesin would be sufficient to perform cohesion, while the remaining fraction would be required for non-cohesive functions. It shows, hypothetically, how a modest decrease in cohesin levels (red dashed line) can sensitize a specific tissue to abnormal development at a given time if cohesin requirement are no longer met. Inspired from Schuster *et al.* 2015.

In the graph, cohesin requirements of cell type 1 peaks sooner than for cell types 2 and 3, and may therefore be the first to suffer from cohesin depletion. Cohesin-SA2 requirements in cardiac progenitors could peak sooner than for other tissues that

undergo differentiation at later times. This model could explain why, at this given point of development, cardiac development is particularly dependent on cohesin-SA2. This is consistent with why there is a much larger transcriptional deregulation in the heart than in the neural tube. The lethality derived from cardiac defects may mask later dependency of other tissue types on cohesin-SA2.

A study has shown that cohesin removal in differentiated macrophages causes limited effects on steady-state transcription but causes great alteration of inducible gene expression during inflammatory response (Cuartero *et al.* 2018). The authors conclude that cohesin's role in transcriptional regulation is most relevant in when the cells needs to change its transcriptional program in response to stimuli. Not unlike this, we can envision how an abrupt rewiring of transcriptional programs needed for the transition of proliferation to lineage-commitment can create an increased requirement of cohesin, most likely of cohesin-SA2.

6. Implications for human disease

Our results in mouse embryos support a causative contribution of cohesin-SA2 function to the congenital heart anomalies detected in CdLS patients, most of them carrying mutations in NIPBL (Chatfield *et al.* 2012). As discussed in the introduction, cohesinopathy cases with *STAG2* mutations have recently been reported too. Among those, male patients carry missense variants and show milder phenotypes that do not include heart defects, while ventricular septal defects and other heart anomalies have been described in female patients carrying loss of function or missense variants (Aoi *et al.* 2019; Kruszka *et al.* 2019; Lehalle *et al.* 2017; Mullegama *et al.* 2019; Soardi *et al.* 2017; Yuan *et al.* 2019). *STAG2* is an X-linked gene, which can explain why germline inactivating mutations are not tolerated in males while heterozygous females may survive through the selection of cells in which the *wild type* allele is not silenced by the X chromosome inactivation process. Accordingly, our model shows that male *Stag2*-null embryos are lethal, while a fraction of *Stag2* heterozygous females can survive until adulthood. It would be of interest to study these females, regarding cohesin content, X chromosome inactivation and developmental aberrations, to gain more insights into human cases.

We have showed evidence supporting the role of cohesin-SA2 in tissue-specific transcription *in vivo* during murine development. We believe that SA2 deficiency and the

consequent transcriptional deregulation could be a pathogenic mechanism underlying the developmental defects that arise in human developmental syndromes caused by *STAG2* mutations.

In the context of tumorigenesis, *STAG2* harbors a significantly higher frequency of inactivating mutations that any other cohesin subunit or regulator. This can be in part due to the fact that it is encoded in the X chromosome and a single mutation is sufficient for inactivation. In addition, SA2 is non-essential for proliferation as it can be replaced by SA1. So far, few functional studies have been performed using cancer-associated SA2 mutations to understand how its loss promotes tumorigenesis, but knowledge about unique functions of cohesin-SA2 is providing insight into this matter.

Given the presence of cohesin-SA2 at regulatory elements and the inability of cohesin-SA1 to localize at these positions, we propose that tumor cells benefit from deregulation of some key genes in the absence of SA2 that confer a favorable environment for tumor progression. Consistent with this idea, in the hematopoietic system, AML-derived cohesin mutations have been shown to promote malignant transformation by altering the balance between self-renewal and differentiation (Mazumdar *et al.* 2015). More specifically, a similar effect was shown upon downregulation of SA2 in hematopoietic precursors. In addition, aberrant expression of lineage commitment genes was directly linked to SA2-only cohesin positions (Viny *et al.* 2019). Therefore, it is possible that effects of total cohesin downregulation could majorly represent the effects of loss of cohesin-SA2.

As mentioned before, both cohesin variants are essential for embryonic development but either one is sufficient for cohesion-related functions and cell viability. Not surprisingly, i*n vitro* studies have recently nominated SA1 as a synthetic lethal target in SA2 mutant cell lines (Benedetti *et al.* 2017; Liu *et al.* 2018; van der Lelij *et al.* 2017). Selective inhibition of cohesin-SA1 has therapeutic potential to treat *STAG2*-mutated tumors, but given that SA1 and SA2 are so similar it will be challenging to find highly specific inhibitors. To this end, structural information about the floppy ends of the SA1 and SA2 proteins might again prove useful. It will be also important to address the toxicity of STAG1 depletion in adult tissues.

On a different note, it has been shown that the aberrant self-renewal and impaired differentiation induced by SA2 ablation in hematopoietic progenitors can be reversed by reactivation of key cohesin-SA2 target genes (Viny *et al.* 2019). This suggests that

restoring expression levels of critical targets regulated by SA2 could have a tumor suppressive effect and could therefore provide novel therapeutic opportunities. Overall, increasing knowledge about the specificities of cohesin-SA1 and cohesin-SA2 can help to further explore these roads.

Conclusions
Conclusiones

Conclusions

1. We successfully generated a conditional knock out mouse model for *Stag2*.

2. SA2 is not essential for cell viability *in vitro*, but MEFs lacking SA2 display reduced proliferation, loosened centromeric cohesion and reduced fidelity in chromosome segregation.

3. Genome-wide distribution of cohesin-SA1 and cohesin-SA2 is not equivalent. Both cohesin variants are found at CTCF sites, but only cohesin-SA2 can occupy non-CTCF cohesin positions enriched in active regulatory elements. Cohesin-SA1 cannot bind to these positions even in the absence of SA2.

4. Association of cohesin-SA2 with chromatin is more dynamic than association of cohesin-SA1. This is likely due to its preferential interaction with cohesin release factor WAPL.

5. Depletion of SA1 or SA2 causes distinct changes on gene expression.

6. Both cohesin variants are essential to fulfil mouse embryonic development. SA2-null embryos are lethal by stage E10.5.

7. SA2-deficient embryos present a widespread developmental delay that correlates with a decrease in cell proliferation and an increase in apoptosis.

8. SA2-null embryos likely die of cardiac anomalies. They display specific malformations in regions derived from the second heart field and present an incorrect deployment of these progenitors into the heart tube at stage E9.5.

9. SA2 ablation causes misexpression of second heart field regulators and alteration of tissue-specific transcription in mouse embryos prior to lethality.

Conclusiones

1. Hemos generado un modelo de ratón mutante condicional para *Stag2*.

2. SA2 no es esencial para la viabilidad celular *in vitro,* pero MEFs deficientes en SA2 presentan una proliferación reducida, una menor cohesion centromérica y una menor fidelidad en segregación cromosómica.

3. La distribución de la cohesina-SA1 y la cohesina-SA2 a lo largo del genoma no es equivalente. Ambas se localizan en sitios de unión a CTCF, pero únicamente la cohesina-SA2 puede ocupar sitios sin CTCF enriquecidos en elementos reguladores de la transcripción. La cohesina-SA1 no puede ubicarse en estas posiciones incluso en ausencia de SA2.

4. La asociación de la cohesina-SA2 a cromatina es más dinámica que la asociación de la cohesina-SA1, probablemente debido a su interacción preferencial con el factor de descarga WAPL.

5. Depleción de SA1 o SA2 produce cambios diferentes en la expresión génica.

6. Ambas variantes de cohesina son esenciales para completar el desarrollo murino. Los embriones deficientes en SA2 son letales a estadío E10.5.

7. Los embriones deficientes en SA2 presentan un retraso generalizado en el desarrollo que correlaciona con una proliferación celular disminuida y un aumento en apoptosis.

8. Es probable que la causa de la letalidad de los embriones deficientes en SA2 sean sus anomalías cardiacas. Desarrollan malformaciones específicas en regiones derivadas del campo secundario y presentan una migración incorrecta de estos progenitores hacia el tubo cardiaco a estadío E9.5.

9. La depleción de SA2 provoca una expresión aberrante de genes reguladores del campo secundario y una alteración en la transcripción específica de tejido en embriones antes de su letalidad.

References

References

Alvarez, S., Diáz, M., Flach, J., … Méndez, J. (2015). Replication stress caused by low MCM expression limits fetal erythropoiesis and hematopoietic stem cell functionality. *Nature Communications*, **6**. doi:10.1038/ncomms9548

Anderson, D. E., Losada, A., Erickson, H. P., & Hirano, T. (2002). Condensin and cohesin display different arm conformations with characteristic hinge angles. *Journal of Cell Biology*, **156**(3), 419–424.

Aoi, H., Lei, M., Mizuguchi, T., … Matsumoto, N. (2019). Nonsense variants in STAG2 result in distinct sex-dependent phenotypes. *Journal of Human Genetics*, **64**(5), 487–492.

Arumugam, P., Nishino, T., Haering, C. H., Gruber, S., & Nasmyth, K. (2006). Cohesin's ATPase activity is stimulated by the C-Terminal winged-helix domain of its kleisin subunit. *Current Biology*, **16**, 1998–2008.

Aziz, K., Sieben, C. J., Jeganathan, K. B., … Van Deursen, J. M. (2018). Mosaic-variegated aneuploidy syndrome mutation or haploinsufficiency in Cep57 impairs tumor suppression. *Journal of Clinical Investigation*, **128**(8), 3517–3534.

Bailey, M. L., Neil, N. J. O., Pel, D. M. Van, Solomon, D. A., Waldman, T., & Hieter, P. (2014). Glioblastoma cells containing mutations in the cohesin component STAG2 are sensitive to PARP inhibition. *Molecular Cancer Therapeutics*, **13**(3), 724–733.

Balbas-Martinez, C., Sagrera, A., Carrillo-de-Santa-Pau, E., … Real, F. X. (2013). Recurrent inactivation of STAG2 in bladder cancer is not associated with aneuploidy. *Nat Genet*, **45**(12), 1464–1469.

Barber, T. D., Mcmanus, K., Yuen, K. W. Y., … Hieter, P. (2008). Chromatid cohesion defects may underlie chromosome instability in human colorectal cancers. *PNAS*, **105**(9), 3443–3448.

Bell, A. C., & Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature*, **405**(6785), 482–485.

Belteki, G., Haigh, J., Kabacs, N., … Nagy, A. (2005). Conditional and inducible transgene expression in mice through the combinatorial use of Cre-mediated recombination and tetracycline induction. *Nucleic Acids Research*, **33**(5), e51.

Benedetti, L., Cereda, M., Monteverde, L. A., Desai, N., & Ciccarelli, F. D. (2017). Synthetic lethal interaction between the tumour suppressor STAG2 and its paralog STAG1. *Oncotarget*, **8**(23), 37619–37632.

Bisht, K. K., Daniloski, Z., & Smith, S. (2013). SA1 binds directly to DNA through its unique AT-hook to promote sister chromatid cohesion at telomeres. *Journal of Cell Science*, **126**(15), 3493–3503.

Biswas, U., Hempel, K., Llano, E., Pendas, A., & Jessberger, R. (2016). Distinct Roles of Meiosis-Specific Cohesin Complexes in Mammalian Spermatogenesis. *PLoS Genetics*, **12**(10), e1006389.

Bonnet, C., Leheup, B., Philippe, C., Gr, M., & Jonveaux, P. (2009). Aberrant GRIA3 transcripts with multi-exon duplications in a family with X-linked mental retardation. *American Journal of Medical Genetics*, **149A**, 1280–1289.

Bruneau, B. G. (2013). Signaling and transcriptional networks in heart development

and regeneration. *Cold Spring Harbor Perspectives in Biology*, **5**. doi:10.1101/cshperspect.a008292

Busslinger, G. A., Stocsits, R. R., Van Der Lelij, P., … Peters, J. M. (2017). Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. *Nature*, **544**(7651), 503–507.

Canudas, S., & Smith, S. (2009). Differential regulation of telomere and centromere cohesion by the Scc3 homologues SA1 and SA2, respectively, in human cells. *Journal of Cell Biology*, **187**(2), 165–173.

Carretero, M., Ruiz-Torres, M., Rodríguez-Corsino, M., Barthelemy, I., & Losada, A. (2013). Pds5B is required for cohesion establishment and Aurora B accumulation at centromeres. *EMBO Journal*, **32**(22), 2938–2949.

Carvalhal, S., Tavares, A., Santos, M. B., Mirkovic, M., & Oliveira, R. A. (2018). A quantitative analysis of cohesin decay in mitotic fidelity. *The Journal of Cell Biology*, **217**(10), 3343–3353.

Chatfield, K. C., Schrier, S. A., Li, J., … Krantz, I. D. (2012). Congenital heart disease in Cornelia de Lange syndrome: Phenotype and genotype analysis. *American Journal of Medical Genetics*, **158A**(10), 2499–2505.

Cuadrado, A., Giménez-Llorente, D., Kojic, A., … Losada, A. (2019). Specific contributions of cohesin-SA1 and cohesin-SA2 to TADs and polycomb domains in embryonic stem cells. *Cell Reports*, **27**, 3500–3510.

Cuadrado, A., Remeseiro, S., Gómez-López, G., Pisano, D. G., & Losada, A. (2012). The specific contributions of cohesin-SA1 to cohesion and gene expression. *Cell Cycle*, **11**(12), 2233–2238.

Cuartero, S., Weiss, F. D., Dharmalingam, G., … Merkenschlager, M. (2018). Control of inducible gene expression links cohesin to hematopoietic progenitor self-renewal and differentiation. *Nature Immunology*, **19**, 932–941.

Davidson, I. F., Bauer, B., Goetz, D., Tang, W., Wutz, G., & Peters, J.-M. (2019). DNA loop extrusion by human cohesin. *Science*, **3418**, 1–13.

de Wit, E., Vos, E. S. M., Holwerda, S. J. B., … de Laat, W. (2015). CTCF Binding Polarity Determines Chromatin Looping. *Molecular Cell*, **60**(4), 676–684.

Deardorff, M. A., Bando, M., Nakato, R., … Shirahige, K. (2012). HDAC8 mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle. *Nature*, **489**(7415), 313–317.

Di Benedetto, D., Musumeci, S. A., Avola, E., … Fichera, M. (2014). Definition of minimal duplicated region encompassing the *XIAP* and *STAG2* genes in the Xq25 microduplication syndrome. *American Journal of Medical Genetics Part A*, **164**(8), 1923–1930.

Epstein, J. A., Aghajanian, H., & Singh, M. K. (2015, February 3). Semaphorin signaling in cardiovascular development. *Cell Metabolism*, Cell Press, pp. 163–173.

Faure, A. J., Schmidt, D., Watt, S., … Flicek, P. (2012). Cohesin regulates tissue-specific expression by stabilizing highly occupied cis-regulatory modules. *Genome Research*, **22**(11), 2163–2175.

Francou, A., Saint-Michel, E., Mesbah, K., … Kelly, R. G. (2013). Second heart field cardiac progenitor cells in the early mouse embryo. *Biochimica et Biophysica Acta*

- *Molecular Cell Research*, **1833**(4), 795–798.

Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., & Mirny, L. A. (2016). Formation of Chromosomal Domains by Loop Extrusion. *Cell Reports*, **15**(9), 2038–2049.

Fujita, Y., Masuda, K., Bando, M., … Yamashita, T. (2017). Decreased cohesin in the brain leads to defective synapse development and anxiety-related behavior. *Journal of Experimental Medicine*, **214**(5), 1431–1452.

Gandhi, R., Gillespie, P. J., & Hirano, T. (2006). Human Wapl is a cohesin-binding protein that promotes sister-chromatid resolution in mitotic prophase. *Current Biology*, **16**(24), 2406–2417.

Gerlich, D., Koch, B., Dupeux, F., Peters, J. M., & Ellenberg, J. (2006). Live-Cell Imaging Reveals a Stable Cohesin-Chromatin Interaction after but Not before DNA Replication. *Current Biology*, **16**(15), 1571–1578.

Gibcus, J. H., & Dekker, J. (2013, March 7). The Hierarchy of the 3D Genome. *Molecular Cell*, pp. 773–782.

Gilbert, S. (2000). Paraxial Mesoderm: The Somites and Their Derivatives. In *Developmental Biology*, 6th edn, Sinauer Associates. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK10085/

Gillespie, P. J., & Hirano, T. (2004). Scc2 couples replication licensing to sister chromatid cohesion in Xenopus egg extracts. *Current*, **14**, 1598–1603.

Gligoris, T. G., Scheinost, J. C., Bürmann, F., … Löwe, J. (2014). Closing the cohesin ring: Structure and function of its Smc3-kleisin interface. *Science*, **346**(6212), 963–967.

Graña, O., López-Fernández, H., Fdez-Riverola, F., González Pisano, D., & Glez-Peña, D. (2018). Bicycle: a bioinformatics pipeline to analyze bisulfite sequencing data. *Bioinformatics*, **34**(8), 1414–1415.

Gröschel, S., Sanders, M. A., Hoogenboezem, R., … Delwel, R. (2014). A single oncogenic enhancer rearrangement causes concomitant EVI1 and GATA2 deregulation in Leukemia. *Cell*, **157**(2), 369–381.

Gruber, S., Arumugam, P., Katou, Y., … Nasmyth, K. (2006). Evidence that loading of cohesin onto chromosomes involves opening of its SMC hinge. *Cell*, **127**(3), 523–537.

Gruber, S., Haering, C. H., & Nasmyth, K. (2003). Chromosomal cohesin forms a ring. *Cell*, **112**(6), 765–777.

Guacci, V., Koshland, D., & Strunnikov, A. (1997). A direct link between sister chromatid cohesion and chromosome condensation revealed through the analysis of MCD1 in S. cerevisiae. *Cell*, **91**(1), 47–57.

Haarhuis, J. H. I., van der Weide, R. H., Blomen, V. A., … Rowland, B. D. (2017). The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. *Cell*, **169**(4), 693-707.e14.

Haering, C. H., Farcas, A. M., Arumugam, P., Metson, J., & Nasmyth, K. (2008). The cohesin ring concatenates sister DNA molecules. *Nature*, **454**(7202), 297–301.

Haferlach, T., Nagata, Y., Grossmann, V., … Ogawa, S. (2014). Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. *Leukemia*, **28**(2), 241–247.

Hansen, A. S., Pustova, I., Cattoglio, C., Tjian, R., & Darzacq, X. (2017). CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *ELife*, **6**, 1–33.

Hara, K., Zheng, G., Qu, Q., … Yu, H. (2014). Structure of cohesin subcomplex pinpoints direct shugoshin-Wapl antagonism in centromeric cohesion. *Nature Structural and Molecular Biology*, **21**(10), 864–870.

Hark, A. T., Schoenherr, C. J., Katz, D. J., Ingram, R. S., Levorse, J. M., & Tilghman, S. M. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature*, **405**(6785), 486–489.

Hauf, S., Roitinger, E., Koch, B., Dittrich, C. M., Mechtler, K., & Peters, J.-M. (2005). Dissociation of cohesin from chromosome arms and loss of arm cohesion during early mitosis depends on phosphorylation of SA2. *PLoS Biology*, **3**(3), e69.

Hauf, S., Waizenegger, I. C., & Peters, J. M. (2001). Cohesin cleavage by separase required for anaphase and cytokinesis in human cells. *Science*, **293**(5533), 1320–1323.

Hayashi, S., Lewis, P., Pevny, L., & McMahon, A. P. (2002). Efficient gene modulation in mouse epiblast using a Sox2Cre transgenic mouse strain. *Gene Expression Patterns*, **2**(1–2), 93–97.

Heidinger-Pauli, J. M., Mert, O., Davenport, C., Guacci, V., & Koshland, D. (2010). Report Systematic Reduction of Cohesin Differentially Affects Chromosome Segregation, Condensation, and DNA Repair. *Current Biology*, **20**, 957–963.

Hirano, T. (2005, April 12). Condensins: Organizing and segregating the genome. *Current Biology*, Cell Press, pp. 265–275.

Hug, C. B., & Vaquerizas, J. M. (2018, December 1). The Birth of the 3D Genome during Early Embryonic Development. *Trends in Genetics*, Elsevier Ltd, pp. 903–914.

Huis In't Veld, P. J., Herzog, F., Ladurner, R., … Peters, J. M. (2014). Characterization of a DNA exit gate in the human cohesin ring. *Science*, **346**(6212), 968–972.

Ibarra-Soria, X., Jawaid, W., Pijuan-Sala, B., … Marioni, J. C. (2018). Defining murine organogenesis at single-cell resolution reveals a role for the leukotriene pathway in regulating blood progenitor formation. *Nature Cell Biology*, **20**(2), 127–134.

Ibn-Salem, J., Köhler, S., Love, M. I., … Robinson, P. N. (2014). Deletions of chromosomal regulatory boundaries are associated with congenital disease. *Genome Biology*, **15**(9), 423.

Izumi, K., Nakato, R., Zhang, Z., … Krantz, I. D. (2015). Germline gain-of-function mutations in AFF4 cause a developmental syndrome functionally linking the super elongation complex and cohesin. *Nature Genetics*, **47**(4), 338–344.

Jan, M., Snyder, T. M., Corces-Zimmerman, M. R., … Majeti, R. (2012). Clonal evolution of preleukemic hematopoietic stem cells precedes human acute myeloid leukemia. *Science Translational Medicine*, **4**(149), 118.

Kagey, M. H., Newman, J. J., Bilodeau, S., … Young, R. A. (2010). Mediator and cohesin connect gene expression and chromatin architecture. *Nature*, **467**(7314), 430–435.

Kandoth, C., McLellan, M. D., Vandin, F., … Ding, L. (2014). Mutational landscape and significance across 12 major cancer types. *Nature*, **502**(7471), 333–339.

Kawauchi, S., Calof, A. L., Santos, R., … Lander, A. D. (2009). Multiple Organ System

Defects and Transcriptional Dysregulation in the Nipbl+/− Mouse, a Model of Cornelia de Lange Syndrome. *PLoS Genetics*, **5**(9), e1000650.

Kawauchi, S., Santos, R., Muto, A., … Calof, A. L. (2016). Using mouse and zebrafish models to understand the etiology of developmental defects in Cornelia de Lange Syndrome. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics*, **172**(2), 138–145.

Kelly, R. G., Buckingham, M. E., & Moorman, A. F. (2014). Heart fields and cardiac morphogenesis. *Cold Spring Harbor Perspectives in Medicine*, **4**(10), 1–10.

Kihara, R., Nagata, Y., Kiyoi, H., … Naoe, T. (2014). Comprehensive analysis of genetic alterations and their prognostic impacts in adult acute myeloid leukemia patients. *Leukemia*, **28**(8), 1586–1595.

Kim, J. S., He, X., Orr, B., … Waldman, T. (2016). Intact Cohesion, Anaphase, and Chromosome Segregation in Human Cells Harboring Tumor-Derived Mutations in STAG2. *PLoS Genetics*, **12**(2), 1–18.

Kojic, A., Cuadrado, A., de Koninck, M., … Losada, A. (2018). Distinct roles of cohesin-SA1 and cohesin-SA2 in 3D chromosome organization. *Nature Structural and Molecular Biology*, **25**, 496–504.

Kon, A., Shih, L. Y., Minamino, M., … Ogawa, S. (2013). Recurrent mutations in multiple components of the cohesin complex in myeloid neoplasms. *Nature Genetics*, **45**(10), 1232–1237.

Kong, X., Ball, A. R., Pham, H. X., … Yokomori, K. (2013). Distinct Functions of Human Cohesin-SA1 and Cohesin-SA2 in Double-Strand Break Repair. *Molecular and Cellular Biology*, **34**(4), 685–698.

Kruszka, P., Berger, S. I., Casa, V., … Muenke, M. (2019). Cohesin complex-associated holoprosencephaly. *Brain: A Journal of Neurology*, **142**(9), 2631–2643.

Kueng, S., Hegemann, B., Peters, B. H., … Peters, J. M. (2006). Wapl Controls the Dynamic Association of Cohesin with Chromatin. *Cell*, **127**(5), 955–967.

Kumar, R., Corbett, M. A., Van Bon, B. W. M., … Gecz, J. (2015). Increased STAG2 dosage defines a novel cohesinopathy with intellectual disability and behavioral problems. *Human Molecular Genetics*, **24**(25), 7171–7181.

Ladurner, R., Kreidl, E., Ivanov, M. P., … Peters, J. (2016). Sororin actively maintains sister chromatid cohesion. *The EMBO Journal*, **35**(6), 635–653.

Lafont, A. L., Song, J., & Rankin, S. (2010). Sororin cooperates with the acetyltransferase Eco2 to ensure DNA replication-dependent sister chromatid cohesion. *Proceedings of the National Academy of Sciences of the United States of America*, **107**(47), 20364–20369.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, **9**(4), 357–359.

Lawrence, M. S., Stojanov, P., Mermel, C. H., … Getz, G. (2014). Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature*, **505**(7484), 495–501.

Lehalle, D., Mosca-Boidron, A. L., Begtrup, A., … Faivre, L. (2017). STAG1 mutations cause a novel cohesinopathy characterised by unspecific syndromic intellectual disability. *Journal of Medical Genetics*, **54**(7), 479–488.

Leiserson, M. D. M., Vandin, F., Wu, H., … Raphael, B. J. (2015). Pan-cancer network

analysis identifies combinations of rare somatic mutations across pathways and protein complexes. *Nature Genetics*, **47**(2), 106–116.

Leroy, C., Jacquemont, M. L., Doray, B., … Malan, V. (2016). Xq25 duplication: The crucial role of the STAG2 gene in this novel human cohesinopathy. *Clinical Genetics*, **89**(1), 68–73.

Lin, Z., Luo, X., & Yu, H. (2016). Structural basis of cohesin cleavage by separase. *Nature*, **532**(7597), 131–134.

Liu, H., Jia, L., & Yu, H. (2013). Phospho-H2A and cohesin specify distinct tension-regulated sgo1 pools at kinetochores and inner centromeres. *Current Biology*, **23**(19), 1927–1933.

Liu, Y., Ji, G., Lu, X., … Zhang, X. (2018). Somatic mutation of the cohesin complex subunit confers therapeutic vulnerabilities in cancer Find the latest version : Somatic mutation of the cohesin complex subunit confers therapeutic vulnerabilities in cancer. *The Journal of Clinical Investigation*, **128**(7), 2951–2965.

Llano, E., Herrán, Y., García-Tuñón, I., … Pendás, A. M. (2012). Meiotic cohesin complexes are essential for the formation of the axial element in mice. *The Journal of Cell Biology*, **197**(7), 877–885.

Losada, A., Hirano, M., & Hirano, T. (1998). Identification of Xenopus SMC protein complexes required for sister chromatid cohesion. *Genes and Development*, **12**(13), 1986–1997.

Losada, A., Hirano, M., & Hirano, T. (2002). Cohesin release is required for sister chromatid resolution, but not for condensin-mediated compaction, at the onset of mitosis. *Genes and Development*, **16**(23), 3004–3016.

Losada, A., Yokochi, T., & Hirano, T. (2005). Functional contribution of Pds5 to cohesin-mediated cohesion in human cells and Xenopus egg extracts. *Journal of Cell Science*, **118**(10), 2133–2141.

Losada, A., Yokochi, T., Kobayashi, R., & Hirano, T. (2000). Identification and characterization of SA/Scc3p subunits in the Xenopus and human cohesin complexes. *Journal of Cell Biology*, **150**(3), 405–416.

Lupiáñez, D. G., Kraft, K., Heinrich, V., … Mundlos, S. (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell*, **161**(5), 1012–1025.

Lupiáñez, D. G., Spielmann, M., & Mundlos, S. (2016, April 1). Breaking TADs: how alterations of chromatin domains result in disease. *Trends in Genetics*, Elsevier Ltd, pp. 225–237.

Martincorena, I., Raine, K. M., Gerstung, M., … Campbell, P. J. (2017). Universal Patterns of Selection in Cancer and Somatic Tissues. *Cell*, **171**, 1029–1041.

Mazumdar, C., Shen, Y., Xavy, S., … Majeti, R. (2015). Leukemia-associated cohesin mutants dominantly enforce stem cell programs and impair human hematopoietic progenitor differentiation. *Cell Stem Cell*, **17**, 1–14.

Mclellan, J. L., Neil, N. J. O., Barrett, I., … Hieter, P. (2012). Synthetic lethality of cohesins with PARPs and replication fork mediators. *PLoS Genetics*, **8**(3). doi:10.1371/journal.pgen.1002574

Meisenberg, C., Pinder, S. I., Hopkins, S. R., … Downs, J. A. (2019). Repression of transcription at DNA Breaks requires cohesin throughout interphase and prevents

genome instability. *Molecular Cell*, **73**(2), 212-223.e7.

Méndez, J., & Stillman, B. (2000). Chromatin Association of Human Origin Recognition Complex, Cdc6, and Minichromosome Maintenance Proteins during the Cell Cycle: Assembly of Prereplication Complexes in Late Mitosis. *Molecular and Cellular Biology*, **20**(22), 8602–8612.

Mi, H., Muruganujan, A., Casagrande, J. T., & Thomas, P. D. (2013). Large-scale gene function analysis with the panther classification system. *Nature Protocols*, **8**(8), 1551–1566.

Michaelis, C., Ciosk, R., & Nasmyth, K. (1997). Cohesins: Chromosomal proteins that prevent premature separation of sister chromatids. *Cell*, **91**(1), 35–45.

Minamino, M., Ishibashi, M., Nakato, R., … Shirahige, K. (2015). Esco1 Acetylates Cohesin via a Mechanism Different from That of Esco2. *Current Biology*, **25**(13), 1694–1706.

Mondal, G., Stevers, M., Goode, B., Ashworth, A., & Solomon, D. A. (2019). A requirement for STAG2 in replication fork progression creates a targetable synthetic lethality in cohesin-mutant cancers. *Nature Communications*, **10**(1), 1686.

Morales, C., Ruiz-Torres, M., Rodríguez-Acebes, S., … Losada, A. (2019). PDS5 proteins are required for proper cohesin dynamics and participate in replication fork protection. *Journal of Biological Chemistry*, jbc-RA119.

Mullegama, S. V., Klein, S. D., Signer, R. H., Vilain, E., & Martinez-Agosto, J. A. (2019). Mutations in STAG2 cause an X-linked cohesinopathy associated with undergrowth, developmental delay, and dysmorphia: Expanding the phenotype in males. *Molecular Genetics and Genomic Medicine*, **7**(2), 1–6.

Mullegama, S. V, Klein, S. D., Mulatinho, M. V, … Martinez-Agosto, J. A. (2017). De novo loss-of-function variants in STAG2 are associated with developmental delay , microcephaly , and congenital anomalies. *American Journal of Medical Genetics*, **173**(5), 1319–1327.

Mullenders, J., Aranda-Orgilles, B., Lhoumaud, P., … Aifantis, I. (2015). Cohesin loss alters adult hematopoietic stem cell homeostasis, leading to myeloproliferative neoplasms. *Journal of Experimental Medicine*, **212**(11), 1833–1850.

Murayama, Y., & Uhlmann, F. (2015). DNA entry into and exit out of the cohesin ring by an interlocking gate mechanism. *Cell*, **163**(7), 1628–1640.

Musacchio, A., & Salmon, E. D. (2007, May). The spindle-assembly checkpoint in space and time. *Nature Reviews Molecular Cell Biology*, pp. 379–393.

Muto, A., Calof, A. L., Lander, A. D., & Schilling, T. F. (2011). Multifactorial origins of heart and gut defects in nipbl-deficient zebrafish, a model of Cornelia de Lange Syndrome. *PLoS Biol*, **9**(10), e1001181.

Nasmyth, K., & Haering, C. H. (2009). Cohesin: Its roles and mechanisms. *Annual Review of Genetics*, **43**(1), 525–558.

Nishiyama, T. (2019). Cohesion and cohesin-dependent chromatin organization. *Current Opinion in Cell Biology*, **58**, 8–14.

Nishiyama, T., Ladurner, R., Schmitz, J., … Peters, J. M. (2010). Sororin mediates sister chromatid cohesion by antagonizing Wapl. *Cell*, **143**(5), 737–749.

Nishiyama, T., Sykora, M. M., Huis, P. J., Mechtler, K., & Peters, J. M. (2013). Aurora B

and Cdk1 mediate Wapl activation and release of acetylated cohesin from chromosomes by phosphorylating Sororin. *Proceedings of the National Academy of Sciences of the United States of America*, **110**(33), 13404–13409.

O'Neil, N. J., van Pel, D. M., & Hieter, P. (2013, May). Synthetic lethality and cancer: Cohesin and PARP at the replication fork. *Trends in Genetics*, pp. 290–297.

Ouyang, Z., Zheng, G., Tomchick, D. R., Luo, X., & Yu, H. (2016). Structural basis and IP6 requirement for Pds5-dependent cohesin dynamics. *Molecular Cell*, **62**(2), 248–259.

Parelho, V., Hadjur, S., Spivakov, M., … Merkenschlager, M. (2008). Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell*, **132**(3), 422–433.

Perez-Garcia, V., Fineberg, E., Wilson, R., … Hemberger, M. (2018). Placentation defects are highly prevalent in embryonic lethal mouse mutants. *Nature*, **555**(7697), 463–468.

Philippe, A., Malan, V., Jacquemont, M.-L., … Cormier-Daire, V. (2013). Xq25 duplications encompassing GRIA3 and STAG2 genes in two families convey recognizable X-linked intellectual disability with distinctive facial appearance. *American Journal of Medical Genetics Part A*, **161**(6), 1370–1375.

Phillips-Cremins, J. E., Sauria, M. E. G., Sanyal, A., … Corces, V. G. (2013). Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell*, **153**(6), 1281–1295.

Piché, J., Van Vliet, P. P., Pucéat, M., & Andelfinger, G. (2019). The expanding phenotypes of cohesinopathies: one ring to rule them all! *Cell Cycle*, **18**(21), 2828–2848.

Ramírez, F., Ryan, D. P., Grüning, B., … Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Research*, **44**(1), 160–165.

Rao, S. S. P., Huang, S. C., Glenn St Hilaire, B., … Aiden, E. L. (2017). Cohesin loss eliminates all loop domains. *Cell*, **171**, 305–320.

Remeseiro, S., Cuadrado, A., Carretero, M., … Losada, A. (2012a). Cohesin-SA1 deficiency drives aneuploidy and tumourigenesis in mice due to impaired replication of telomeres. *EMBO Journal*, **31**(9), 2076–2089.

Remeseiro, S., Cuadrado, A., Gómez-Lãpez, G., Pisano, D. G., & Losada, A. (2012b). A unique role of cohesin-SA1 in gene regulation and development. *EMBO Journal*, **31**(9), 2090–2102.

Remeseiro, S., Cuadrado, A., Kawauchi, S., Calof, A. L., Lander, A. D., & Losada, A. (2013a). Reduction of Nipbl impairs cohesin loading locally and affects transcription but not cohesion-dependent functions in a mouse model of Cornelia de Lange Syndrome. *Biochimica et Biophysica Acta - Molecular Basis of Disease*, **1832**(12), 2097–2102.

Remeseiro, S., Cuadrado, A., & Losada, A. (2013b). Cohesin in development and disease. *Development*, **140**(18), 3715–3718.

Roig, M. B., Löwe, J., Chan, K. L., Beckouët, F., Metson, J., & Nasmyth, K. (2014). Structure and function of cohesin's Scc3/SA regulatory subunit. *FEBS Letters*, **588**(20), 3692–3702.

Rubio, E. D., Reiss, D. J., Welcsh, P. L., … Krumm, A. (2008). CTCF physically links cohesin to chromatin. *Proceedings of the National Academy of Sciences of the United States of America*, **105**(24), 8309–8314.

Ruzankina, Y., Pinzon-Guzman, C., Asare, A., … Brown, E. J. (2007). Deletion of the developmentally essential gene ATR in adult mice leads to age-related phenotypes and stem cell loss. *Cell Stem Cell*, **1**(1), 113–126.

Sanborn, A. L., Rao, S. S. P., Huang, S. C., … Aiden, E. L. (2015). Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proceedings of the National Academy of Sciences of the United States of America*, **112**(47), E6456–E6465.

Santos, R., Kawauchi, S., Jacobs, R. E., … Calof, A. L. (2016). Conditional Creation and Rescue of Nipbl-Deficiency in Mice Reveals Multiple Determinants of Risk for Congenital Heart Defects. *PLOS Biology*, **14**(9), e2000197.

Schmidt, D., Schwalie, P. C., Ross-Innes, C. S., … Odom, D. T. (2010). A CTCF-independent role for cohesin in tissue-specific transcription. *Genome Research*, **20**(5), 578–588.

Schuster, K., Leeke, B., Meier, M., … Horsfield, J. A. (2015). A neural crest origin for cohesinopathy heart defects. *Human Molecular Genetics*, ddv402.

Schwarzer, W., Abdennur, N., Goloborodko, A., … Spitz, F. (2017). Two independent modes of chromatin organization revealed by cohesin removal. *Nature*, **551**(7678), 51–56.

Scialdone, A., Tanaka, Y., Jawaid, W., … Göttgens, B. (2016). Resolving early mesoderm diversification through single-cell expression profiling. *Nature*, **535**(7611), 289–293.

Shintomi, K., & Hirano, T. (2009). Releasing cohesin from chromosome arms in early mitosis: Opposing actions of Wapl-Pds5 and Sgo1. *Genes and Development*, **23**(18), 2224–2236.

Soardi, F. C., Machado-Silva, A., Linhares, N. D., … Pena, S. D. J. (2017). Familial Stag2 germline mutation defines a new human cohesinopathy. *Genomic Medicine*, **2**(1), 7.

Sofueva, S., Yaffe, E., Chan, W. C., … Hadjur, S. (2013). Cohesin-mediated interactions organize chromosomal domain architecture. *EMBO Journal*, **32**(24), 3119–3129.

Solomon, D. A., Kim, T., Diaz-martinez, L. A., … Waldman, T. (2011). Mutational inactivation of STAG2 causes aneuploidy in human cancer. *Science*, **333**(6045), 1039–1043.

Subramanian, A., Kuehn, H., Gould, J., Tamayo, P., & Mesirov, J. P. (2007). GSEA-P: a desktop application for Gene Set Enrichment Analysis. *Bioinformatics*, **23**(23), 3251–3253.

Sumara, I., Vorlaufer, E., Gieffers, C., Peters, B. H., & Peters, J. M. (2000). Characterization of vertebrate cohesin complexes and their regulation in prophase. *Journal of Cell Biology*, **151**(4), 749–761.

Tedeschi, A., Wutz, G., Huet, S., … Peters, J. M. (2013). Wapl is an essential regulator of chromatin structure and chromosome segregation. *Nature*, **501**(7468), 564–568.

Thol, F., Bollin, R., Gehlhaar, M., … Heuser, M. (2014). Mutations in the cohesin complex in acute myeloid leukemia: Clinical and prognostic implications. *Blood*, **123**(6), 914–920.

Thota, S., Viny, A. D., Makishima, H., … Maciejewski, J. P. (2014). Genetic alterations of the cohesin complex genes in myeloid malignancies. *Blood*, **124**(11), 1790–1798.

Trevino, A. E., & Zhang, F. (2014). Genome editing using cas9 nickases. In *Methods in Enzymology*, Vol. 546, Academic Press Inc., pp. 161–174.

Uhlmann, F. (2016, July 1). SMC complexes: From DNA to chromosomes. *Nature Reviews Molecular Cell Biology*, Nature Publishing Group, pp. 399–412.

Uhlmann, F., Lottspelch, F., & Nasmyth, K. (1999). Sister-chromatid separation at anaphase onset is promoted by cleavage of the cohesin subunit Scc1. *Nature*, **400**(6739), 37–42.

van der Lelij, P., Lieb, S., Jude, J., … Petronczki, M. (2017). Synthetic lethality between the cohesin subunits STAG1 and STAG2 in diverse cancer contexts. *ELife*, **6**, 1–15.

Van Der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, **9**, 2579–2625.

Viny, A. D., Bowman, R. L., Liu, Y., … Levine, R. L. (2019). Cohesin members Stag1 and Stag2 display distinct roles in chromatin accessibility and topological control of HSC self-renewal and differentiation. *Cell Stem Cell*, **25**(5), 682-696.e8.

Viny, A. D., Ott, C. J., Spitzer, B., … Levine, R. L. (2015). Dose-dependent role of the cohesin complex in normal and malignant hematopoiesis. *Journal of Experimental Medicine*, **212**(11), 1819–1832.

Waizenegger, I. C., Hauf, S., Meinke, A., & Peters, J. M. (2000). Two distinct pathways remove mammalian cohesin from chromosome arms in prophase and from centromeres in anaphase. *Cell*, **103**(3), 399–410.

Wang, T., Glover, B., Hadwiger, G., Miller, C. A., di Martino, O., & Welch, J. S. (2019). Smc3 is required for mouse embryonic and adult hematopoiesis. *Experimental Hematology*, **70**, 70–84.

Welch, J. S., Ley, T. J., Link, D. C., … Wilson, R. K. (2012). The origin and evolution of mutations in acute myeloid leukemia. *Cell*, **150**, 264–278.

Wendt, K. S., Yoshida, K., Itoh, T., … Peters, J. M. (2008a). Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature*, **451**(7180), 796–801.

Wendt, K. S., Yoshida, K., Itoh, T., … Peters, J. M. (2008b). Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature*, **451**(7180), 796–801.

Whelan, G., Kreidl, E., Wutz, G., Egner, A., Peters, J.-M., & Eichele, G. (2012). Cohesin acetyltransferase Esco2 is a cell viability factor and is required for cohesion in pericentric heterochromatin. *The EMBO Journal*, **31**(1), 71–82.

White, J. K., Gerdin, A. K., Karp, N. A., … Steel, K. P. (2013). Genome-wide generation and systematic phenotyping of knockout mice reveals new roles for many genes. *Cell*, **154**(2), 452.

Wutz, G., Várnai, C., Nagasaka, K., … Peters, J. (2017). Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. *The EMBO Journal*, **36**(24), 3573–3599.

Xiao, T., Wallace, J., & Felsenfeld, G. (2011). Specific sites in the C terminus of CTCF interact with the SA2 subunit of the Cohesin complex and are required for Cohesin-dependent insulation activity. *Molecular and Cellular Biology*, **31**(11), 2174–2183.

Xu, H., Balakrishnan, K., Malaterre, J., … McKay, M. J. (2010). Rad21-Cohesin Haploinsufficiency Impedes DNA Repair and Enhances Gastrointestinal Radiosensitivity in Mice. *PLoS ONE*, **5**(8), e12112.

Yingjun, X., Wen, T., Yujian, L., … Junhong, C. (2015). Microduplication of chromosome Xq25 encompassing STAG2 gene in a boy with intellectual disability. *European Journal of Medical Genetics*, **58**(2), 116–121.

Yuan, B., Neira, J., Pehlivan, D., … Liu, P. (2019). Clinical exome sequencing reveals locus heterogeneity and phenotypic variability of cohesinopathies. *Genetics in Medicine*, **21**(3), 663–675.

Zhang, N., Kuznetsov, S. G., Sharan, S. K., Li, K., Rao, P. H., & Pati, D. (2008a). A handcuff model for the cohesin complex. *Journal of Cell Biology*, **183**(6), 1019–1031.

Zhang, Y., Liu, T., Meyer, C. A., … Shirley, X. S. (2008b). Model-based analysis of ChIP-Seq (MACS). *Genome Biology*, **9**(9). doi:10.1186/gb-2008-9-9-r137

Zheng, H., & Xie, W. (2019, September 1). The role of 3D genome organization in development and cell differentiation. *Nature Reviews Molecular Cell Biology*, Nature Publishing Group, pp. 535–550.

Zuin, J., Dixon, J. R., Van Der Reijden, M. I. J. A., … Wendt, K. S. (2014). Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proceedings of the National Academy of Sciences of the United States of America*, **111**(3), 996–1001.

Alvarez, S., Diáz, M., Flach, J., … Méndez, J. (2015). Replication stress caused by low MCM expression limits fetal erythropoiesis and hematopoietic stem cell functionality. *Nature Communications*, **6**. doi:10.1038/ncomms9548

Anderson, D. E., Losada, A., Erickson, H. P., & Hirano, T. (2002). Condensin and cohesin display different arm conformations with characteristic hinge angles. *Journal of Cell Biology*, **156**(3), 419–424.

Aoi, H., Lei, M., Mizuguchi, T., … Matsumoto, N. (2019). Nonsense variants in STAG2 result in distinct sex-dependent phenotypes. *Journal of Human Genetics*, **64**(5), 487–492.

Arumugam, P., Nishino, T., Haering, C. H., Gruber, S., & Nasmyth, K. (2006). Cohesin's ATPase activity is stimulated by the C-Terminal winged-helix domain of its kleisin subunit. *Current Biology*, **16**, 1998–2008.

Aziz, K., Sieben, C. J., Jeganathan, K. B., … Van Deursen, J. M. (2018). Mosaic-variegated aneuploidy syndrome mutation or haploinsufficiency in Cep57 impairs tumor suppression. *Journal of Clinical Investigation*, **128**(8), 3517–3534.

Bailey, M. L., Neil, N. J. O., Pel, D. M. Van, Solomon, D. A., Waldman, T., & Hieter, P. (2014). Glioblastoma cells containing mutations in the cohesin component STAG2 are sensitive to PARP inhibition. *Molecular Cancer Therapeutics*, **13**(3), 724–733.

Balbas-Martinez, C., Sagrera, A., Carrillo-de-Santa-Pau, E., … Real, F. X. (2013). Recurrent inactivation of STAG2 in bladder cancer is not associated with aneuploidy. *Nat Genet*, **45**(12), 1464–1469.

Barber, T. D., Mcmanus, K., Yuen, K. W. Y., … Hieter, P. (2008). Chromatid cohesion defects may underlie chromosome instability in human colorectal cancers. *PNAS*, **105**(9), 3443–3448.

Bell, A. C., & Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature*, **405**(6785), 482–485.

Belteki, G., Haigh, J., Kabacs, N., … Nagy, A. (2005). Conditional and inducible transgene expression in mice through the combinatorial use of Cre-mediated recombination and tetracycline induction. *Nucleic Acids Research*, **33**(5), e51.

Benedetti, L., Cereda, M., Monteverde, L. A., Desai, N., & Ciccarelli, F. D. (2017). Synthetic lethal interaction between the tumour suppressor STAG2 and its paralog STAG1. *Oncotarget*, **8**(23), 37619–37632.

Bisht, K. K., Daniloski, Z., & Smith, S. (2013). SA1 binds directly to DNA through its unique AT-hook to promote sister chromatid cohesion at telomeres. *Journal of Cell Science*, **126**(15), 3493–3503.

Biswas, U., Hempel, K., Llano, E., Pendas, A., & Jessberger, R. (2016). Distinct Roles of Meiosis-Specific Cohesin Complexes in Mammalian Spermatogenesis. *PLoS Genetics*, **12**(10), e1006389.

Bonnet, C., Leheup, B., Philippe, C., Gr, M., & Jonveaux, P. (2009). Aberrant GRIA3 transcripts with multi-exon duplications in a family with X-linked mental retardation. *American Journal of Medical Genetics*, **149A**, 1280–1289.

Bruneau, B. G. (2013). Signaling and transcriptional networks in heart development and regeneration. *Cold Spring Harbor Perspectives in Biology*, **5**. doi:10.1101/cshperspect.a008292

Busslinger, G. A., Stocsits, R. R., Van Der Lelij, P., … Peters, J. M. (2017). Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. *Nature*, **544**(7651), 503–507.

Canudas, S., & Smith, S. (2009). Differential regulation of telomere and centromere cohesion by the Scc3 homologues SA1 and SA2, respectively, in human cells. *Journal of Cell Biology*, **187**(2), 165–173.

Carretero, M., Ruiz-Torres, M., Rodríguez-Corsino, M., Barthelemy, I., & Losada, A. (2013). Pds5B is required for cohesion establishment and Aurora B accumulation at centromeres. *EMBO Journal*, **32**(22), 2938–2949.

Carvalhal, S., Tavares, A., Santos, M. B., Mirkovic, M., & Oliveira, R. A. (2018). A quantitative analysis of cohesin decay in mitotic fidelity. *The Journal of Cell Biology*, **217**(10), 3343–3353.

Chatfield, K. C., Schrier, S. A., Li, J., … Krantz, I. D. (2012). Congenital heart disease in Cornelia de Lange syndrome: Phenotype and genotype analysis. *American Journal of Medical Genetics*, **158A**(10), 2499–2505.

Cuadrado, A., Giménez-Llorente, D., Kojic, A., … Losada, A. (2019). Specific contributions of cohesin-SA1 and cohesin-SA2 to TADs and polycomb domains in embryonic stem cells. *Cell Reports*, **27**, 3500–3510.

Cuadrado, A., Remeseiro, S., Gómez-López, G., Pisano, D. G., & Losada, A. (2012). The specific contributions of cohesin-SA1 to cohesion and gene expression. *Cell Cycle*, **11**(12), 2233–2238.

Cuartero, S., Weiss, F. D., Dharmalingam, G., … Merkenschlager, M. (2018). Control

of inducible gene expression links cohesin to hematopoietic progenitor self-renewal and differentiation. *Nature Immunology*, **19**, 932–941.

Davidson, I. F., Bauer, B., Goetz, D., Tang, W., Wutz, G., & Peters, J.-M. (2019). DNA loop extrusion by human cohesin. *Science*, **3418**, 1–13.

de Wit, E., Vos, E. S. M., Holwerda, S. J. B., … de Laat, W. (2015). CTCF Binding Polarity Determines Chromatin Looping. *Molecular Cell*, **60**(4), 676–684.

Deardorff, M. A., Bando, M., Nakato, R., … Shirahige, K. (2012). HDAC8 mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle. *Nature*, **489**(7415), 313–317.

Di Benedetto, D., Musumeci, S. A., Avola, E., … Fichera, M. (2014). Definition of minimal duplicated region encompassing the *XIAP* and *STAG2* genes in the Xq25 microduplication syndrome. *American Journal of Medical Genetics Part A*, **164**(8), 1923–1930.

Epstein, J. A., Aghajanian, H., & Singh, M. K. (2015, February 3). Semaphorin signaling in cardiovascular development. *Cell Metabolism*, Cell Press, pp. 163–173.

Faure, A. J., Schmidt, D., Watt, S., … Flicek, P. (2012). Cohesin regulates tissue-specific expression by stabilizing highly occupied cis-regulatory modules. *Genome Research*, **22**(11), 2163–2175.

Francou, A., Saint-Michel, E., Mesbah, K., … Kelly, R. G. (2013). Second heart field cardiac progenitor cells in the early mouse embryo. *Biochimica et Biophysica Acta - Molecular Cell Research*, **1833**(4), 795–798.

Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., & Mirny, L. A. (2016). Formation of Chromosomal Domains by Loop Extrusion. *Cell Reports*, **15**(9), 2038–2049.

Fujita, Y., Masuda, K., Bando, M., … Yamashita, T. (2017). Decreased cohesin in the brain leads to defective synapse development and anxiety-related behavior. *Journal of Experimental Medicine*, **214**(5), 1431–1452.

Gandhi, R., Gillespie, P. J., & Hirano, T. (2006). Human Wapl is a cohesin-binding protein that promotes sister-chromatid resolution in mitotic prophase. *Current Biology*, **16**(24), 2406–2417.

Gerlich, D., Koch, B., Dupeux, F., Peters, J. M., & Ellenberg, J. (2006). Live-Cell Imaging Reveals a Stable Cohesin-Chromatin Interaction after but Not before DNA Replication. *Current Biology*, **16**(15), 1571–1578.

Gibcus, J. H., & Dekker, J. (2013, March 7). The Hierarchy of the 3D Genome. *Molecular Cell*, pp. 773–782.

Gilbert, S. (2000). Paraxial Mesoderm: The Somites and Their Derivatives. In *Developmental Biology*, 6th edn, Sinauer Associates. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK10085/

Gillespie, P. J., & Hirano, T. (2004). Scc2 couples replication licensing to sister chromatid cohesion in Xenopus egg extracts. *Current*, **14**, 1598–1603.

Gligoris, T. G., Scheinost, J. C., Bürmann, F., … Löwe, J. (2014). Closing the cohesin ring: Structure and function of its Smc3-kleisin interface. *Science*, **346**(6212), 963–967.

Graña, O., López-Fernández, H., Fdez-Riverola, F., González Pisano, D., & Glez-

Peña, D. (2018). Bicycle: a bioinformatics pipeline to analyze bisulfite sequencing data. *Bioinformatics*, **34**(8), 1414–1415.

Gröschel, S., Sanders, M. A., Hoogenboezem, R., … Delwel, R. (2014). A single oncogenic enhancer rearrangement causes concomitant EVI1 and GATA2 deregulation in Leukemia. *Cell*, **157**(2), 369–381.

Gruber, S., Arumugam, P., Katou, Y., … Nasmyth, K. (2006). Evidence that loading of cohesin onto chromosomes involves opening of its SMC hinge. *Cell*, **127**(3), 523–537.

Gruber, S., Haering, C. H., & Nasmyth, K. (2003). Chromosomal cohesin forms a ring. *Cell*, **112**(6), 765–777.

Guacci, V., Koshland, D., & Strunnikov, A. (1997). A direct link between sister chromatid cohesion and chromosome condensation revealed through the analysis of MCD1 in S. cerevisiae. *Cell*, **91**(1), 47–57.

Haarhuis, J. H. I., van der Weide, R. H., Blomen, V. A., … Rowland, B. D. (2017). The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. *Cell*, **169**(4), 693-707.e14.

Haering, C. H., Farcas, A. M., Arumugam, P., Metson, J., & Nasmyth, K. (2008). The cohesin ring concatenates sister DNA molecules. *Nature*, **454**(7202), 297–301.

Haferlach, T., Nagata, Y., Grossmann, V., … Ogawa, S. (2014). Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. *Leukemia*, **28**(2), 241–247.

Hansen, A. S., Pustova, I., Cattoglio, C., Tjian, R., & Darzacq, X. (2017). CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *ELife*, **6**, 1–33.

Hara, K., Zheng, G., Qu, Q., … Yu, H. (2014). Structure of cohesin subcomplex pinpoints direct shugoshin-Wapl antagonism in centromeric cohesion. *Nature Structural and Molecular Biology*, **21**(10), 864–870.

Hark, A. T., Schoenherr, C. J., Katz, D. J., Ingram, R. S., Levorse, J. M., & Tilghman, S. M. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature*, **405**(6785), 486–489.

Hauf, S., Roitinger, E., Koch, B., Dittrich, C. M., Mechtler, K., & Peters, J.-M. (2005). Dissociation of cohesin from chromosome arms and loss of arm cohesion during early mitosis depends on phosphorylation of SA2. *PLoS Biology*, **3**(3), e69.

Hauf, S., Waizenegger, I. C., & Peters, J. M. (2001). Cohesin cleavage by separase required for anaphase and cytokinesis in human cells. *Science*, **293**(5533), 1320–1323.

Hayashi, S., Lewis, P., Pevny, L., & McMahon, A. P. (2002). Efficient gene modulation in mouse epiblast using a Sox2Cre transgenic mouse strain. *Gene Expression Patterns*, **2**(1–2), 93–97.

Heidinger-Pauli, J. M., Mert, O., Davenport, C., Guacci, V., & Koshland, D. (2010). Report Systematic Reduction of Cohesin Differentially Affects Chromosome Segregation, Condensation, and DNA Repair. *Current Biology*, **20**, 957–963.

Hirano, T. (2005, April 12). Condensins: Organizing and segregating the genome. *Current Biology*, Cell Press, pp. 265–275.

Hug, C. B., & Vaquerizas, J. M. (2018, December 1). The Birth of the 3D Genome during Early Embryonic Development. *Trends in Genetics*, Elsevier Ltd, pp. 903–

914.

Huis In't Veld, P. J., Herzog, F., Ladurner, R., … Peters, J. M. (2014). Characterization of a DNA exit gate in the human cohesin ring. *Science*, **346**(6212), 968–972.

Ibarra-Soria, X., Jawaid, W., Pijuan-Sala, B., … Marioni, J. C. (2018). Defining murine organogenesis at single-cell resolution reveals a role for the leukotriene pathway in regulating blood progenitor formation. *Nature Cell Biology*, **20**(2), 127–134.

Ibn-Salem, J., Köhler, S., Love, M. I., … Robinson, P. N. (2014). Deletions of chromosomal regulatory boundaries are associated with congenital disease. *Genome Biology*, **15**(9), 423.

Izumi, K., Nakato, R., Zhang, Z., … Krantz, I. D. (2015). Germline gain-of-function mutations in AFF4 cause a developmental syndrome functionally linking the super elongation complex and cohesin. *Nature Genetics*, **47**(4), 338–344.

Jan, M., Snyder, T. M., Corces-Zimmerman, M. R., … Majeti, R. (2012). Clonal evolution of preleukemic hematopoietic stem cells precedes human acute myeloid leukemia. *Science Translational Medicine*, **4**(149), 118.

Kagey, M. H., Newman, J. J., Bilodeau, S., … Young, R. A. (2010). Mediator and cohesin connect gene expression and chromatin architecture. *Nature*, **467**(7314), 430–435.

Kandoth, C., McLellan, M. D., Vandin, F., … Ding, L. (2014). Mutational landscape and significance across 12 major cancer types. *Nature*, **502**(7471), 333–339.

Kawauchi, S., Calof, A. L., Santos, R., … Lander, A. D. (2009). Multiple Organ System Defects and Transcriptional Dysregulation in the Nipbl+/− Mouse, a Model of Cornelia de Lange Syndrome. *PLoS Genetics*, **5**(9), e1000650.

Kawauchi, S., Santos, R., Muto, A., … Calof, A. L. (2016). Using mouse and zebrafish models to understand the etiology of developmental defects in Cornelia de Lange Syndrome. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics*, **172**(2), 138–145.

Kelly, R. G., Buckingham, M. E., & Moorman, A. F. (2014). Heart fields and cardiac morphogenesis. *Cold Spring Harbor Perspectives in Medicine*, **4**(10), 1–10.

Kihara, R., Nagata, Y., Kiyoi, H., … Naoe, T. (2014). Comprehensive analysis of genetic alterations and their prognostic impacts in adult acute myeloid leukemia patients. *Leukemia*, **28**(8), 1586–1595.

Kim, J. S., He, X., Orr, B., … Waldman, T. (2016). Intact Cohesion, Anaphase, and Chromosome Segregation in Human Cells Harboring Tumor-Derived Mutations in STAG2. *PLoS Genetics*, **12**(2), 1–18.

Kojic, A., Cuadrado, A., de Koninck, M., … Losada, A. (2018). Distinct roles of cohesin-SA1 and cohesin-SA2 in 3D chromosome organization. *Nature Structural and Molecular Biology*, **25**, 496–504.

Kon, A., Shih, L. Y., Minamino, M., … Ogawa, S. (2013). Recurrent mutations in multiple components of the cohesin complex in myeloid neoplasms. *Nature Genetics*, **45**(10), 1232–1237.

Kong, X., Ball, A. R., Pham, H. X., … Yokomori, K. (2013). Distinct Functions of Human Cohesin-SA1 and Cohesin-SA2 in Double-Strand Break Repair. *Molecular and Cellular Biology*, **34**(4), 685–698.

Kruszka, P., Berger, S. I., Casa, V., … Muenke, M. (2019). Cohesin complex-

associated holoprosencephaly. *Brain: A Journal of Neurology*, **142**(9), 2631–2643.

Kueng, S., Hegemann, B., Peters, B. H., … Peters, J. M. (2006). Wapl Controls the Dynamic Association of Cohesin with Chromatin. *Cell*, **127**(5), 955–967.

Kumar, R., Corbett, M. A., Van Bon, B. W. M., … Gecz, J. (2015). Increased STAG2 dosage defines a novel cohesinopathy with intellectual disability and behavioral problems. *Human Molecular Genetics*, **24**(25), 7171–7181.

Ladurner, R., Kreidl, E., Ivanov, M. P., … Peters, J. (2016). Sororin actively maintains sister chromatid cohesion. *The EMBO Journal*, **35**(6), 635–653.

Lafont, A. L., Song, J., & Rankin, S. (2010). Sororin cooperates with the acetyltransferase Eco2 to ensure DNA replication-dependent sister chromatid cohesion. *Proceedings of the National Academy of Sciences of the United States of America*, **107**(47), 20364–20369.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, **9**(4), 357–359.

Lawrence, M. S., Stojanov, P., Mermel, C. H., … Getz, G. (2014). Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature*, **505**(7484), 495–501.

Lehalle, D., Mosca-Boidron, A. L., Begtrup, A., … Faivre, L. (2017). STAG1 mutations cause a novel cohesinopathy characterised by unspecific syndromic intellectual disability. *Journal of Medical Genetics*, **54**(7), 479–488.

Leiserson, M. D. M., Vandin, F., Wu, H., … Raphael, B. J. (2015). Pan-cancer network analysis identifies combinations of rare somatic mutations across pathways and protein complexes. *Nature Genetics*, **47**(2), 106–116.

Leroy, C., Jacquemont, M. L., Doray, B., … Malan, V. (2016). Xq25 duplication: The crucial role of the STAG2 gene in this novel human cohesinopathy. *Clinical Genetics*, **89**(1), 68–73.

Lin, Z., Luo, X., & Yu, H. (2016). Structural basis of cohesin cleavage by separase. *Nature*, **532**(7597), 131–134.

Liu, H., Jia, L., & Yu, H. (2013). Phospho-H2A and cohesin specify distinct tension-regulated sgo1 pools at kinetochores and inner centromeres. *Current Biology*, **23**(19), 1927–1933.

Liu, Y., Ji, G., Lu, X., … Zhang, X. (2018). Somatic mutation of the cohesin complex subunit confers therapeutic vulnerabilities in cancer Find the latest version : Somatic mutation of the cohesin complex subunit confers therapeutic vulnerabilities in cancer. *The Journal of Clinical Investigation*, **128**(7), 2951–2965.

Llano, E., Herrán, Y., García-Tuñón, I., … Pendás, A. M. (2012). Meiotic cohesin complexes are essential for the formation of the axial element in mice. *The Journal of Cell Biology*, **197**(7), 877–885.

Losada, A., Hirano, M., & Hirano, T. (1998). Identification of Xenopus SMC protein complexes required for sister chromatid cohesion. *Genes and Development*, **12**(13), 1986–1997.

Losada, A., Hirano, M., & Hirano, T. (2002). Cohesin release is required for sister chromatid resolution, but not for condensin-mediated compaction, at the onset of mitosis. *Genes and Development*, **16**(23), 3004–3016.

Losada, A., Yokochi, T., & Hirano, T. (2005). Functional contribution of Pds5 to

cohesin-mediated cohesion in human cells and Xenopus egg extracts. *Journal of Cell Science*, **118**(10), 2133–2141.

Losada, A., Yokochi, T., Kobayashi, R., & Hirano, T. (2000). Identification and characterization of SA/Scc3p subunits in the Xenopus and human cohesin complexes. *Journal of Cell Biology*, **150**(3), 405–416.

Lupiáñez, D. G., Kraft, K., Heinrich, V., … Mundlos, S. (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell*, **161**(5), 1012–1025.

Lupiáñez, D. G., Spielmann, M., & Mundlos, S. (2016, April 1). Breaking TADs: how alterations of chromatin domains result in disease. *Trends in Genetics*, Elsevier Ltd, pp. 225–237.

Martincorena, I., Raine, K. M., Gerstung, M., … Campbell, P. J. (2017). Universal Patterns of Selection in Cancer and Somatic Tissues. *Cell*, **171**, 1029–1041.

Mazumdar, C., Shen, Y., Xavy, S., … Majeti, R. (2015). Leukemia-associated cohesin mutants dominantly enforce stem cell programs and impair human hematopoietic progenitor differentiation. *Cell Stem Cell*, **17**, 1–14.

Mclellan, J. L., Neil, N. J. O., Barrett, I., … Hieter, P. (2012). Synthetic lethality of cohesins with PARPs and replication fork mediators. *PLoS Genetics*, **8**(3). doi:10.1371/journal.pgen.1002574

Meisenberg, C., Pinder, S. I., Hopkins, S. R., … Downs, J. A. (2019). Repression of transcription at DNA Breaks requires cohesin throughout interphase and prevents genome instability. *Molecular Cell*, **73**(2), 212-223.e7.

Méndez, J., & Stillman, B. (2000). Chromatin Association of Human Origin Recognition Complex, Cdc6, and Minichromosome Maintenance Proteins during the Cell Cycle: Assembly of Prereplication Complexes in Late Mitosis. *Molecular and Cellular Biology*, **20**(22), 8602–8612.

Mi, H., Muruganujan, A., Casagrande, J. T., & Thomas, P. D. (2013). Large-scale gene function analysis with the panther classification system. *Nature Protocols*, **8**(8), 1551–1566.

Michaelis, C., Ciosk, R., & Nasmyth, K. (1997). Cohesins: Chromosomal proteins that prevent premature separation of sister chromatids. *Cell*, **91**(1), 35–45.

Minamino, M., Ishibashi, M., Nakato, R., … Shirahige, K. (2015). Esco1 Acetylates Cohesin via a Mechanism Different from That of Esco2. *Current Biology*, **25**(13), 1694–1706.

Mondal, G., Stevers, M., Goode, B., Ashworth, A., & Solomon, D. A. (2019). A requirement for STAG2 in replication fork progression creates a targetable synthetic lethality in cohesin-mutant cancers. *Nature Communications*, **10**(1), 1686.

Morales, C., Ruiz-Torres, M., Rodríguez-Acebes, S., … Losada, A. (2019). PDS5 proteins are required for proper cohesin dynamics and participate in replication fork protection. *Journal of Biological Chemistry*, jbc-RA119.

Mullegama, S. V., Klein, S. D., Signer, R. H., Vilain, E., & Martinez-Agosto, J. A. (2019). Mutations in STAG2 cause an X-linked cohesinopathy associated with undergrowth, developmental delay, and dysmorphia: Expanding the phenotype in males. *Molecular Genetics and Genomic Medicine*, **7**(2), 1–6.

Mullegama, S. V, Klein, S. D., Mulatinho, M. V, … Martinez-Agosto, J. A. (2017). De novo loss-of-function variants in STAG2 are associated with developmental delay , microcephaly , and congenital anomalies. *American Journal of Medical Genetics*, **173**(5), 1319–1327.

Mullenders, J., Aranda-Orgilles, B., Lhoumaud, P., … Aifantis, I. (2015). Cohesin loss alters adult hematopoietic stem cell homeostasis, leading to myeloproliferative neoplasms. *Journal of Experimental Medicine*, **212**(11), 1833–1850.

Murayama, Y., & Uhlmann, F. (2015). DNA entry into and exit out of the cohesin ring by an interlocking gate mechanism. *Cell*, **163**(7), 1628–1640.

Musacchio, A., & Salmon, E. D. (2007, May). The spindle-assembly checkpoint in space and time. *Nature Reviews Molecular Cell Biology*, pp. 379–393.

Muto, A., Calof, A. L., Lander, A. D., & Schilling, T. F. (2011). Multifactorial origins of heart and gut defects in nipbl-deficient zebrafish, a model of Cornelia de Lange Syndrome. *PLoS Biol*, **9**(10), e1001181.

Nasmyth, K., & Haering, C. H. (2009). Cohesin: Its roles and mechanisms. *Annual Review of Genetics*, **43**(1), 525–558.

Nishiyama, T. (2019). Cohesion and cohesin-dependent chromatin organization. *Current Opinion in Cell Biology*, **58**, 8–14.

Nishiyama, T., Ladurner, R., Schmitz, J., … Peters, J. M. (2010). Sororin mediates sister chromatid cohesion by antagonizing Wapl. *Cell*, **143**(5), 737–749.

Nishiyama, T., Sykora, M. M., Huis, P. J., Mechtler, K., & Peters, J. M. (2013). Aurora B and Cdk1 mediate Wapl activation and release of acetylated cohesin from chromosomes by phosphorylating Sororin. *Proceedings of the National Academy of Sciences of the United States of America*, **110**(33), 13404–13409.

O'Neil, N. J., van Pel, D. M., & Hieter, P. (2013, May). Synthetic lethality and cancer: Cohesin and PARP at the replication fork. *Trends in Genetics*, pp. 290–297.

Ouyang, Z., Zheng, G., Tomchick, D. R., Luo, X., & Yu, H. (2016). Structural basis and IP6 requirement for Pds5-dependent cohesin dynamics. *Molecular Cell*, **62**(2), 248–259.

Parelho, V., Hadjur, S., Spivakov, M., … Merkenschlager, M. (2008). Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell*, **132**(3), 422–433.

Perez-Garcia, V., Fineberg, E., Wilson, R., … Hemberger, M. (2018). Placentation defects are highly prevalent in embryonic lethal mouse mutants. *Nature*, **555**(7697), 463–468.

Philippe, A., Malan, V., Jacquemont, M.-L., … Cormier-Daire, V. (2013). Xq25 duplications encompassing GRIA3 and STAG2 genes in two families convey recognizable X-linked intellectual disability with distinctive facial appearance. *American Journal of Medical Genetics Part A*, **161**(6), 1370–1375.

Phillips-Cremins, J. E., Sauria, M. E. G., Sanyal, A., … Corces, V. G. (2013). Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell*, **153**(6), 1281–1295.

Piché, J., Van Vliet, P. P., Pucéat, M., & Andelfinger, G. (2019). The expanding phenotypes of cohesinopathies: one ring to rule them all! *Cell Cycle*, **18**(21), 2828–2848.

Ramírez, F., Ryan, D. P., Grüning, B., … Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Research*, **44**(1), 160–165.

Rao, S. S. P., Huang, S. C., Glenn St Hilaire, B., … Aiden, E. L. (2017). Cohesin loss eliminates all loop domains. *Cell*, **171**, 305–320.

Remeseiro, S., Cuadrado, A., Carretero, M., … Losada, A. (2012a). Cohesin-SA1 deficiency drives aneuploidy and tumourigenesis in mice due to impaired replication of telomeres. *EMBO Journal*, **31**(9), 2076–2089.

Remeseiro, S., Cuadrado, A., Gómez-Lãpez, G., Pisano, D. G., & Losada, A. (2012b). A unique role of cohesin-SA1 in gene regulation and development. *EMBO Journal*, **31**(9), 2090–2102.

Remeseiro, S., Cuadrado, A., Kawauchi, S., Calof, A. L., Lander, A. D., & Losada, A. (2013a). Reduction of Nipbl impairs cohesin loading locally and affects transcription but not cohesion-dependent functions in a mouse model of Cornelia de Lange Syndrome. *Biochimica et Biophysica Acta - Molecular Basis of Disease*, **1832**(12), 2097–2102.

Remeseiro, S., Cuadrado, A., & Losada, A. (2013b). Cohesin in development and disease. *Development*, **140**(18), 3715–3718.

Roig, M. B., Löwe, J., Chan, K. L., Beckouët, F., Metson, J., & Nasmyth, K. (2014). Structure and function of cohesin's Scc3/SA regulatory subunit. *FEBS Letters*, **588**(20), 3692–3702.

Rubio, E. D., Reiss, D. J., Welcsh, P. L., … Krumm, A. (2008). CTCF physically links cohesin to chromatin. *Proceedings of the National Academy of Sciences of the United States of America*, **105**(24), 8309–8314.

Ruzankina, Y., Pinzon-Guzman, C., Asare, A., … Brown, E. J. (2007). Deletion of the developmentally essential gene ATR in adult mice leads to age-related phenotypes and stem cell loss. *Cell Stem Cell*, **1**(1), 113–126.

Sanborn, A. L., Rao, S. S. P., Huang, S. C., … Aiden, E. L. (2015). Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proceedings of the National Academy of Sciences of the United States of America*, **112**(47), E6456–E6465.

Santos, R., Kawauchi, S., Jacobs, R. E., … Calof, A. L. (2016). Conditional Creation and Rescue of Nipbl-Deficiency in Mice Reveals Multiple Determinants of Risk for Congenital Heart Defects. *PLOS Biology*, **14**(9), e2000197.

Schmidt, D., Schwalie, P. C., Ross-Innes, C. S., … Odom, D. T. (2010). A CTCF-independent role for cohesin in tissue-specific transcription. *Genome Research*, **20**(5), 578–588.

Schuster, K., Leeke, B., Meier, M., … Horsfield, J. A. (2015). A neural crest origin for cohesinopathy heart defects. *Human Molecular Genetics*, ddv402.

Schwarzer, W., Abdennur, N., Goloborodko, A., … Spitz, F. (2017). Two independent modes of chromatin organization revealed by cohesin removal. *Nature*, **551**(7678), 51–56.

Scialdone, A., Tanaka, Y., Jawaid, W., … Göttgens, B. (2016). Resolving early mesoderm diversification through single-cell expression profiling. *Nature*, **535**(7611), 289–293.

Shintomi, K., & Hirano, T. (2009). Releasing cohesin from chromosome arms in early mitosis: Opposing actions of Wapl-Pds5 and Sgo1. *Genes and Development*, **23**(18), 2224–2236.

Soardi, F. C., Machado-Silva, A., Linhares, N. D., … Pena, S. D. J. (2017). Familial Stag2 germline mutation defines a new human cohesinopathy. *Genomic Medicine*, **2**(1), 7.

Sofueva, S., Yaffe, E., Chan, W. C., … Hadjur, S. (2013). Cohesin-mediated interactions organize chromosomal domain architecture. *EMBO Journal*, **32**(24), 3119–3129.

Solomon, D. A., Kim, T., Diaz-martinez, L. A., … Waldman, T. (2011). Mutational inactivation of STAG2 causes aneuploidy in human cancer. *Science*, **333**(6045), 1039–1043.

Subramanian, A., Kuehn, H., Gould, J., Tamayo, P., & Mesirov, J. P. (2007). GSEA-P: a desktop application for Gene Set Enrichment Analysis. *Bioinformatics*, **23**(23), 3251–3253.

Sumara, I., Vorlaufer, E., Gieffers, C., Peters, B. H., & Peters, J. M. (2000). Characterization of vertebrate cohesin complexes and their regulation in prophase. *Journal of Cell Biology*, **151**(4), 749–761.

Tedeschi, A., Wutz, G., Huet, S., … Peters, J. M. (2013). Wapl is an essential regulator of chromatin structure and chromosome segregation. *Nature*, **501**(7468), 564–568.

Thol, F., Bollin, R., Gehlhaar, M., … Heuser, M. (2014). Mutations in the cohesin complex in acute myeloid leukemia: Clinical and prognostic implications. *Blood*, **123**(6), 914–920.

Thota, S., Viny, A. D., Makishima, H., … Maciejewski, J. P. (2014). Genetic alterations of the cohesin complex genes in myeloid malignancies. *Blood*, **124**(11), 1790–1798.

Trevino, A. E., & Zhang, F. (2014). Genome editing using cas9 nickases. In *Methods in Enzymology*, Vol. 546, Academic Press Inc., pp. 161–174.

Uhlmann, F. (2016, July 1). SMC complexes: From DNA to chromosomes. *Nature Reviews Molecular Cell Biology*, Nature Publishing Group, pp. 399–412.

Uhlmann, F., Lottspelch, F., & Nasmyth, K. (1999). Sister-chromatid separation at anaphase onset is promoted by cleavage of the cohesin subunit Scc1. *Nature*, **400**(6739), 37–42.

van der Lelij, P., Lieb, S., Jude, J., … Petronczki, M. (2017). Synthetic lethality between the cohesin subunits STAG1 and STAG2 in diverse cancer contexts. *ELife*, **6**, 1–15.

Van Der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, **9**, 2579–2625.

Viny, A. D., Bowman, R. L., Liu, Y., … Levine, R. L. (2019). Cohesin members Stag1 and Stag2 display distinct roles in chromatin accessibility and topological control of HSC self-renewal and differentiation. *Cell Stem Cell*, **25**(5), 682-696.e8.

Viny, A. D., Ott, C. J., Spitzer, B., … Levine, R. L. (2015). Dose-dependent role of the cohesin complex in normal and malignant hematopoiesis. *Journal of Experimental Medicine*, **212**(11), 1819–1832.

Waizenegger, I. C., Hauf, S., Meinke, A., & Peters, J. M. (2000). Two distinct pathways remove mammalian cohesin from chromosome arms in prophase and from centromeres in anaphase. *Cell*, **103**(3), 399–410.

Wang, T., Glover, B., Hadwiger, G., Miller, C. A., di Martino, O., & Welch, J. S. (2019). Smc3 is required for mouse embryonic and adult hematopoiesis. *Experimental Hematology*, **70**, 70–84.

Welch, J. S., Ley, T. J., Link, D. C., … Wilson, R. K. (2012). The origin and evolution of mutations in acute myeloid leukemia. *Cell*, **150**, 264–278.

Wendt, K. S., Yoshida, K., Itoh, T., … Peters, J. M. (2008a). Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature*, **451**(7180), 796–801.

Wendt, K. S., Yoshida, K., Itoh, T., … Peters, J. M. (2008b). Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature*, **451**(7180), 796–801.

Whelan, G., Kreidl, E., Wutz, G., Egner, A., Peters, J.-M., & Eichele, G. (2012). Cohesin acetyltransferase Esco2 is a cell viability factor and is required for cohesion in pericentric heterochromatin. *The EMBO Journal*, **31**(1), 71–82.

White, J. K., Gerdin, A. K., Karp, N. A., … Steel, K. P. (2013). Genome-wide generation and systematic phenotyping of knockout mice reveals new roles for many genes. *Cell*, **154**(2), 452.

Wutz, G., Várnai, C., Nagasaka, K., … Peters, J. (2017). Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. *The EMBO Journal*, **36**(24), 3573–3599.

Xiao, T., Wallace, J., & Felsenfeld, G. (2011). Specific sites in the C terminus of CTCF interact with the SA2 subunit of the Cohesin complex and are required for Cohesin-dependent insulation activity. *Molecular and Cellular Biology*, **31**(11), 2174–2183.

Xu, H., Balakrishnan, K., Malaterre, J., … McKay, M. J. (2010). Rad21-Cohesin Haploinsufficiency Impedes DNA Repair and Enhances Gastrointestinal Radiosensitivity in Mice. *PLoS ONE*, **5**(8), e12112.

Yingjun, X., Wen, T., Yujian, L., … Junhong, C. (2015). Microduplication of chromosome Xq25 encompassing STAG2 gene in a boy with intellectual disability. *European Journal of Medical Genetics*, **58**(2), 116–121.

Yuan, B., Neira, J., Pehlivan, D., … Liu, P. (2019). Clinical exome sequencing reveals locus heterogeneity and phenotypic variability of cohesinopathies. *Genetics in Medicine*, **21**(3), 663–675.

Zhang, N., Kuznetsov, S. G., Sharan, S. K., Li, K., Rao, P. H., & Pati, D. (2008a). A handcuff model for the cohesin complex. *Journal of Cell Biology*, **183**(6), 1019–1031.

Zhang, Y., Liu, T., Meyer, C. A., … Shirley, X. S. (2008b). Model-based analysis of ChIP-Seq (MACS). *Genome Biology*, **9**(9). doi:10.1186/gb-2008-9-9-r137

Zheng, H., & Xie, W. (2019, September 1). The role of 3D genome organization in development and cell differentiation. *Nature Reviews Molecular Cell Biology*, Nature Publishing Group, pp. 535–550.

Zuin, J., Dixon, J. R., Van Der Reijden, M. I. J. A., … Wendt, K. S. (2014). Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proceedings of the National Academy of Sciences of the United States of*

Annex